Improving Document Navigation Using Space-Filling Thumbnails

November 10, 2005

Jason Alexander
jma142@student.canterbury.ac.nz
Department of Computer Science and Software Engineering
University of Canterbury, Christchurch, New Zealand

Supervisor: Assoc. Prof. Dr. Andy Cockburn
andy@cosc.canterbury.ac.nz

Abstract

Space-Filling Thumbnails (SFT) is a new technique for document navigation that disposes of traditional scrollbars. SFT provides a matrix of thumbnail images, one for every page of the document, on a *single* screen. Pages can be selected for viewing by using a mouse.

A prototype Space-Filling Thumbnails interface has been designed, developed and implemented. Six further document navigation systems have been implemented to enable comparative evaluations.

Three formal experiments have been carried out, all of which have shown Space-Filling Thumbnails to be more efficient than other research and mainstream document navigation systems. The first experiment ($n = 13$) showed SFT to be significantly faster than all other document navigation interfaces for both visual search and spatial re-acquisition tasks. Experiment Two ($n = 32$) showed that SFT is robust and faster than Thumbnail-Enhanced Scrollbars (TES) for visual and spatial search tasks in documents of differing lengths. Lastly, in Experiment Three ($n = 12$), the possible "Achilles' Heel" of SFT was tested—long documents. SFT proved to be more efficient than TES even in its "worst-case scenario" situation—a 300 page book. Subjective measures from these experiments showed favour for SFT.

We have shown that Space-Filling Thumbnails outperforms traditional document navigation systems through the exploitation of users' spatial memory. The swiftness of visual and spatial searches in SFT have proved that this system has the potential to revolutionise document navigation, as a valuable addition to, if not replacement for, scrollbars.

Acknowledgements

The author would like to thank his supervisor Andy Cockburn for his inspiration, enthusiasm and guidance throughout this project. He would also like to thank his family, staff members and fellow post-graduate students who have shown interest and offered suppor. All of the participants involved in the experiments also deserve thanks, for giving up their time and providing valuable feedback.

Contents

1 Introduction

Document navigation involves moving through a particular document, or moving among a series of documents. This study focuses on the former—navigation within a single document.

It has long been acknowledged by researchers that the scrollbar is an impediment to efficient navigation within a document. In a taxonomic study of web navigation Bryne et al. [4] found users spend a large proportion of their web-browsing time scrolling. They noted that scrolling is “an obvious case where widget design could make a difference”.

Taking this on board, researchers have attempted to improve document navigation by making small incremental improvements to the traditional scrollbar. The predominant improvement to navigation with scrollbars is the addition of a column of thumbnails, as found in products such as Adobe Reader[1] and recent versions of Microsoft Word[2].

Many of the issues regarding electronic document navigation arise due to the difficulties in maintaining the spatial properties of a physical page when it is transferred to its electronic equivalent. This is often due to the physical restrictions on the size of screens. In their comparison of paper and online document navigation, O’Hara and Sellen [28] reported that participants using electronic navigation talked of the need to be able to “lay pages out in order to form a mental picture or overview of the document,” something that is easily achieved when using the paper based equivalent. They concluded that “laying out pages in space was found to [be] important for gaining an overall sense of the documents”. This ability to give an overview of a document is something that current navigation interfaces lack. Space-Filling Thumbnails provides an overview of the complete document implicitly. Thumbnail images of all pages are laid out in a matrix, giving the user an overview of the structure and length of the document.

In this investigation we argue that efficient document navigation cannot be achieved through a traditional scrollbar. A conceptual design shift is required to optimise search efficiency. Document navigation interfaces should be built *around* navigation tasks, rather than having the functionality included as an afterthought.

We propose Space-Filling Thumbnails (SFT), a promising new technique for document navigation. The Space-Filling Thumbnails interface displays an array of thumbnails, one for each page in the document, on a single screen. Pages can be viewed at normal reading magnification using a simple mouse click. The user can return to the thumbnail view with a further mouse click. This system allows efficient visual search, exploits a user’s spatial memory and allows an overview of the document to be rapidly obtained. The onscreen location of the thumbnails will generally be static, allowing frequently visited pages to be quickly accessed, through spatial re-acquisition.

The idea of viewing thumbnails of a document is not new. Microsoft PowerPoint will display thumbnail images of the slides in a presentation when “Slide Sorter” mode is selected. However, the thumbnails are not arranged in an optimal space-filling manner and often require scrolling to access the complete presentation. Microsoft Word displays all of the pages in a document in “Print Preview” mode (if configured). This mode is not designed to support rapid navigation between thumbnail view and normal view.

Robertson and Mackinlay [31] alluded to the SFT design in their work on the *Document Lens*. “If you lay the entire contents of a multi-page document out in two dimensions so it is all visible, the text will typically be much too small to read”, they go on to say “yet we would like to be able to do this so that patterns in the document can be easily perceived”. Their solution to this was to use a distortion oriented approach where a lens is moved across the layout of pages and “stretched to provide a continuous display in

[1] http://www.adobe.com

[2] http://www.microsoft.com

the global context". The SFT system does not attempt to make pages readable in the thumbnails view—we believe the page structure shown in the thumbnails provides sufficient information for identification.

A prototype Space-Filling Thumbnails interface has been designed and implemented. Six further document navigation systems—Multi-Page Rapid Serial Visual Presentation, Rate-Based Scrolling, Rapid Serial Visual Presentation, Scrollbars, Thumbnail-Enhanced Scrollbars and Speed-Dependent Automatic Zooming have been implemented so as to allow comparative evaluations to take place.

The three formal evaluations presented here show SFT to out-perform all other document navigation systems in both visual search and spatial re-acquisition tasks. Subjective evaluations of the system show it to be favourable with users.

The structure of this report is as follows: Chapter Two describes related work and previously proposed document navigation systems. Chapter Three provides details of the design and implementation of the systems used for evaluation. A description of the comparative studies conducted is provided in Chapter Four. A discussion of this work, limitations and future research is presented in Chapter Five and conclusions in Chapter Six.

2 Related Work

This chapter describes previous work that has been conducted in the area of document navigation. We first describe document navigation interfaces and then move to related areas of spatial memory. Finally, we briefly describe Fitts' and Hick-Hyman Laws and their applicability to interaction with Space-Filling Thumbnails.

2.1 Document Navigation Interfaces

This chapter presents several document navigation interfaces, both research and commercial, that have been proposed.

2.1.1 Scrollbars

The scrollbar has become the standard (and most widely applied) widget for navigation purposes, be it for documents, images, diagrams or any other object displayable electronically. The scrollbar attempts to provide spatial cues as to the current location within the document through the position of the scroll thumb in the scroll trough.

There are several disadvantages of the scrollbar that result in it being an inefficient navigation mechanism. In large documents the scroll thumb quickly becomes inaccurate and clumsy when directly manipulated. This is exemplified when a one pixel movement in the scroll thumb results in several pages in the document scrolling. Further, the scrollbar does not explicitly encourage the use of spatial memory to help in re-acquisition situations. It is difficult to remember the exact location of the scroll thumb in its trough, especially when the thumb reaches its minimum size. Thirdly, Smith and Schraefel [33], noted that "only if users are familiar with the document can they use the scrollbar to quickly locate the approximate position of interest. . . otherwise they are limited to scrolling through the document from top to bottom"—the scrollbar provides us with no clues as to the content of the document.

Many incremental improvements have been proposed to assist the functionality of basic scrollbars. These include visualisations within and around the scrollbar, to aid the user in search tasks.

The bookmark scrollbar [22] allows the user to place bookmarks of interesting document features onto an area beside the scroll-trough. Although similar to bookmark facilities available in applications such as Adobe Reader, this system holds two key advantages. Firstly, the bookmarks are placed on the scroll-trough allowing users to tailor the positions of bookmarks. Secondly, the explicit act of placing the bookmarks inherently helps the user to remember the location of interesting sections.

Byrd [3] developed a visualisation of information *within* a standard scrollbar. Their focus was on aiding the search for particular terms in a document. The text of the document contained each of the search terms highlighted in different colours. The scrollbar then contained a visualisation of these highlighted search terms to give the user an indication as to where in the document each of the terms appears. Effectively, this resulted in small coloured dots in the scrollbar indicating word placements (a "scrollbar with confetti"). This allows users to quickly navigate to positions in the document that are of the most interest [3]. Within scrollbar visualisations have also been proposed by McCrickard and Catrambone [25] that allow the user to view an outline of the structure of the document in the scroll trough.

Perhaps the most widely implemented improvement to scrollbars is the addition of a column of thumbnails. This improvement is described in the following section.

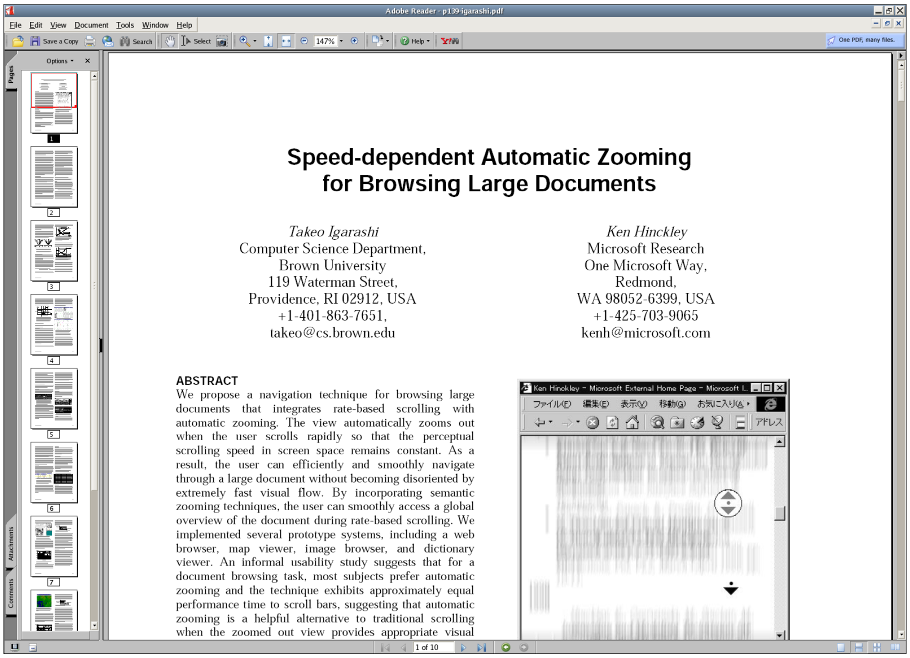


Figure 2.1: The Adobe Reader interface, illustrating Thumbnail-Enhanced Scrollbars

Thumbnail-Enhanced Scrollbars

Thumbnail-Enhanced Scrollbars add two components to the standard scrollbar based navigation. Firstly, they retain the standard scrollbar and viewing pane in majority of the window area. Then in a second panel (typically to the left of the main pane) a series of thumbnails of the pages in the document are presented within another scrolling environment. This setup is typical of popular document navigation interfaces, such as Adobe Reader, as illustrated in figure 2.1.

Researchers have attempted to improve on this idea further, by eliminating the scrolling environment around the thumbnails. The document viewer, *DeckView* described by Ginsburg et al. [13] makes use of a column of overlapping thumbnails. The thumbnail of the current page is always fully-rendered, but other thumbnails are placed overlapping each other to prevent the need for scrollbars. Navigation through the thumbnails is achieved by dragging through the thumbnail trough, with each thumbnail being brought to the front and then returned to its position.

2.1.2 Rate-Based Scrolling (RBS) Techniques

Scrollbar based systems are *position control* systems that rely on the user to move the position of the scroll thumb in some manner. In contrast to this, Rate-Based Scrolling is a *displacement* based system that requires the user to control navigation by displacing the mouse. Rate-based scrolling is commonly found on Microsoft Windows applications, initiated by clicking and dragging with the middle mouse button. The speed at which the document flows past the user is determined by how far the mouse is displaced from the position at which the rate-based scroll was initiated.

Speed Dependent Automatic Zooming (SDAZ)

Speed Dependant Automatic Zooming (SDAZ) is a variation of rate-based scrolling that aims to reduce motion blur while moving rapidly through a document. The idea was originally proposed by Igarashi and Hinckley [19]. The navigation interface “automatically binds a document’s zoom level with its scroll-speed” [32]. This allows the viewer to get a broader overview of a section of a document “more quickly and with less effort than traditional document navigation techniques” [32]. To illustrate, Savage and Cock-

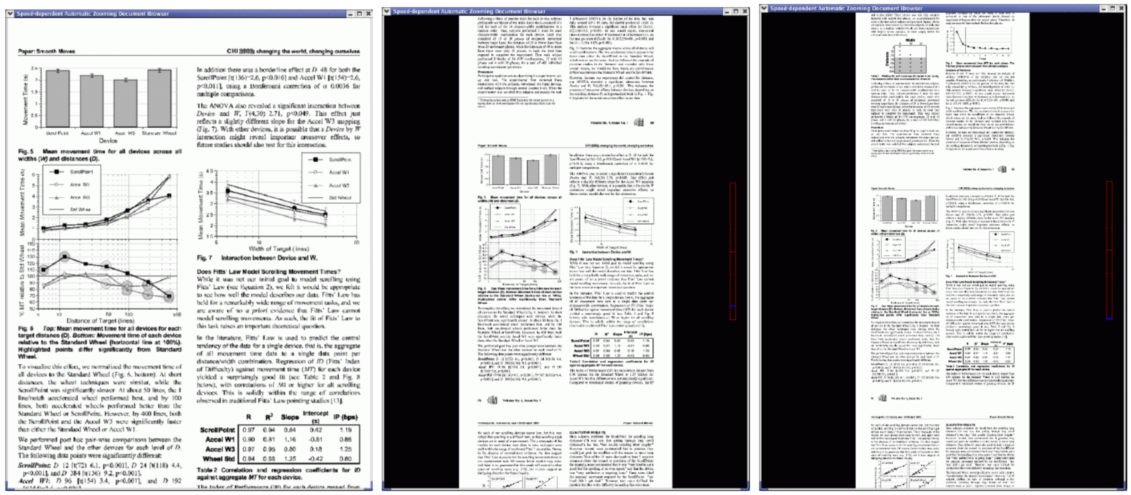


Figure 2.2: Speed Dependant Automatic Zooming at slow, medium and fast scrolling speeds (courtesy [32])

burn produced a time-series of screen-shots, which is reproduced in Figure 2.2. The leftmost image shows the interface scrolling at slow speed, the middle image at medium speed and the rightmost image depicts scrolling at a fast speed [32]. In their evaluation of SDAZ, Cockburn and Savage [6] found it to be significantly faster than other scrolling systems (scrollbars, rate-based scrolling and Displacement-Dependent Automatic Zooming) for visual search tasks. They also found users would prefer to use SDAZ over any of the other systems.

2.1.3 Rapid Serial Visual Presentation (RSVP) Techniques

Rapid Serial Visual Presentation (RSVP) interfaces expose the user to a set of images at a rapid rate (drawn from the technique of flipping through the pages of a book) [34]. RSVP exploits the humans ability to rapidly recognise distinctive images. The feature that distinguishes RSVP from other techniques is that it allows “images to be presented sequentially in the time-domain, threreby offering an alternative to the conventional concurrent display of image[s] in the space domain” [34].

There are two distinctive types of RSVP: static and dynamic. Static RSVP refers to the method of displaying images in the same location. Each image in a static RSVP “stream” is only visible until it is replaced by the next. Dynamic RSVP refers to a method of displaying the images in several locations—an image will be seen in its context. One example of this is rapidly flipping thumbnails of the images around in an arc, allowing each image to be on screen for a longer time [8].

RSVP interfaces can also be classified according to how many images they present simultaneously (note, the namespaces over-lap). Spence et al. [34] described three types of RSVP. “Static” mode RSVP displays *all* of the images simultaneously, similar to the thumbnail arrangement proposed in our Space-Filling Thumbnails interface. “Mixed” mode RSVP involves displaying *four* images simultaneously and “Slide Show” mode involves displaying one image at a time.

The presentation rate of the images in a RSVP system heavily influences its usability. The human is able to recognise images displayed for as little as 100ms, or even less [34]. When the images are displayed for less than this time, images are “momentarily understood at the time of viewing and then quickly forgotten” [8]. This allows the detection of a particular image, but does not aid the user in forming a model of the complete document.

The RSVP technique has been applied to a variety of information visualisation areas, for example TV channel selection, web-browsing and PDA applications [34]. The Flipper system proposed by Sun and Guimbretière [35] appears to be the first system to *explicitly* implement RSVP as the main navigation technique in a *Document Navigation* interface (note that Adobe Reader configured in non-continuous scroll mode can function in an RSVP browser). Most previous evaluations have concentrated on flipping images, rather than document pages. Flipper combines SDAZ with RSVP, to provide a smooth transition from traditional scrolling to a page-flipping mode. Experiments with Flipper found that for visual search tasks it was significantly faster than an SDAZ implementation and a system imitating Adobe Reader [35]. Interestingly they found no significant difference between the SDAZ implementation and their *Scrolling with Thumbnails* condition. This contradicts experimental findings by Cockburn et al. [6, 5], that SDAZ is faster than an Adobe Reader system. Sun and Guimbretière offer little explanation for this, except that the systems were implemented differently. It is likely the use of thumbnails affected their results—Cockburn et al's experiments did not use thumbnails, but the evaluation with Flipper did, showing that there are efficiency gains to be had when using thumbnails to aid search tasks.

Hoeben and Stappers [15] developed a thumbnail-based RSVP navigation system. Instead of the whole page flipping into view, only a small thumbnail was flipped through the pages. The main viewing pane is synchronised when the flipping is halted. No evaluation of this system was performed, but the authors note that their system was not a replacement for scrollbars, but was more suited to search tasks.

2.1.4 Flip Zoom Interfaces

The *Hierarchical Flip Zoom* interface proposed by Holmquist and Björk [2] is a focus+context[1] interface that provides a preview of all pages in a document or image library via thumbnails. Each page in the document is laid on a discrete tile. The tile in focus is substantially larger than the surrounding thumbnails and is placed in the center of the screen. Context is retained by shifting the other tiles into sequential order around the focused page. This idea is extended further by allowing multiple hierarchies of documents to be displayed simultaneously, with each tile now representing a document with the visualisation recursively applied.

2.1.5 PhotoMesa

Although not applied to document navigation, the PhotoMesa interface developed by Bederson [1] is worthy of mention because, like Space-Filling Thumbnails, it presents a series of thumbnail images in a space-filling manner. The thumbnails are arranged into groups using quantum treemaps, according to their directory or some auxiliary meta-data. To view a group of images the user can click on one of the images contained within that group. Further intermediate zooming steps are available until reaching the required viewing level, or, double-clicking on a specific image zooms it to occupy the complete screen.

2.2 Spatial Memory

Human spatial memory is responsible for recording information regarding the individual's environment and spatial orientation. In relation to Graphical User Interfaces (GUIs), spatial memory is the ability to remember where on an interface a particular object is located.

Our ability to remember the location of an object can be strongly influenced by the effort required to locate that object. Ehret's study [10] showed that increasing the cost of obtaining an item improves the chance of remembering the spatial location of that object. This is exploited in interfaces such as the *Data Mountain*, proposed by Robertson et al. [30]. A user is required to place items on the 3D interface where they desire. This is a higher effort task than simple selection and so allows users to remember the location of the object more easily. Their studies showed that this ability to arrange items in a 3-dimensional space improved re-acquisition time. Czerwinski et al. [7] investigated the use of spatial memory along with other memory aiding devices for locating web pages in the data mountain. They concluded that "pictorial image[s], spatial location and mouse-over text title can be quite effective" when searching for objects in space.

Users do not have to consciously try to remember the location of buttons or labels for them to be placed into their spatial memory. Many interfaces present static depictions of information, and do not allow it to be manipulated. In these situations (which are common, for instance the selection of a toolbar button), a user is more interested in completing their task as quickly as possible than being worried about future

[1] A focus+context interface is a system that allows a user to focus on one page, simultaneously maintaining it in the context of the document.

re-acquisition of that target. For this reason Mandler et al's [24] result regarding the processing of location information is important, reporting that we incidentally and effortlessly process location information.

Due to this incidentality Jones and Dumais [21] warn that spatial memory should not be relied on for effective location of objects. Their research shows that combining a spatial location with some form of symbolic representation improves performance. This is the standard for toolbar and menu mechanisms—a label or icon on a button in a constant position.

Devices to aid spatial memory whilst navigating through documents have also been proposed, usually to ease the return to frequently viewed pages. Hoeben and Stappers [15] implemented an "Electronic Dog-ear" system, by displaying dog-earred pages for a long time during an RSVP visualisation. This provides the user with a longer time to react to those pages that are of most interest.

2.2.1 Spatial Memory and Scrolling

O'Hara et al. [29] discuss the implications of scrollbars on spatial memory. They point out that with scrollbars there is a "dynamic relationship between the location of the information and the reference points such as window parameters". The information being displayed to the user is only constant with respect to the pages around it. The current page of focus is always moving with respect to window decorations, toolbars and menus.

The fixity of information in a physical document allows spatial information to be more readily obtained and retained than in electronic documents. This "fixity of information" allows users of paper based media to acquire "incidental knowledge of the location of information by reference to its physical place on the page" [28]. With a piece of paper an image at the top of the page will *always* be at the top of the page. Whereas with any technique that employs scrolling, the image my appear anywhere on the screen—top, middle bottom or not at all. Dumais and Jones [9] re-iterates that "importantly, a paper-based object has an attribute of spatiality…its contents occupy a definite point in three-dimensional space," something not readily available on a restricted 2D screen space.

2.3 Power Law of Practise

In this investigation we are interested in how the target page search time decreases as the same pages are repetitively located. The Power Law of Practice states that task performance time decreases with practice [27]. The biggest gains in performance occur at the start of practice with decreasing gains the longer the practice is conducted. This rate of learning is often referred to as the *learning curve* [26]. This law is now taken as the benchmark of any system that involves learning, such that any system that performs worse than that predicted by this law cannot be taken seriously. The Power Law of Practise can be stated mathematically as:

$$log(T_n) \quad = \quad C - \alpha log(n)$$

where $C = T_1$, the time taken to complete the task on the first trial and α is the learning curve steepness. T_n is the task time after n practise sessions. The α value is of particular note—the larger this value, the quicker learning is taking place.

2.4 Hick-Hyman Law for Choice Reaction

The Space-Filling Thumbnails interface requires the user to chose which of the n pages in a document they wish to visit. Hick-Hyman law [18] governs the reaction time t, (in seconds) a user takes to make a decision between C equally probable options. This is described mathematically as:

$$t \quad = \quad a + b \times log_2(C)$$

Where, a and b are empirically derived constants. The slope b decreases with large amounts of practise. This law generally holds when each of the choices are easily distinguishable [23]. This will be the case in SFT until documents become very large and hence the thumbnails become very small.

2.5 Summary

We have discussed some of the numerous proposed document navigation systems and implications of spatial memory. The de facto standard for navigation remains the scrollbar along with one of the many enhancements, combining it with a column of thumbnails. Several research systems have been discussed, including SDAZ and RSVP techniques, both of which prove to be faster for visual search tasks than traditional scrollbars.

The incidental learning of spatial locations has been discussed, important because the Space-Filling Thumbnails system aims to exploit this memory feature. We also discussed laws that govern both how acquisition time should improve with practise and how long it takes to chose from equally probable options.

3 Design and Implementation

This chapter presents the design and implementation of the document navigation systems used in the formal evaluations. Six previously proposed interfaces were implemented along with Space-Filling Thumbnails.

All of these systems were implemented using the same core C++/OpenGL base to allow fair comparisons to be drawn. The systems chosen for evaluation represent a mix of those commonly used in practise and promising research systems. Details on the evaluations performed can be found in Chapter 4.

A full description of the SFT interface is provided first, along with implementation details of the other systems. The chapter finishes with a description of the control software and hardware apparatus used in the experiments.

3.1 Space-Filling Thumbnails

The Space-Filling Thumbnails system arranges a matrix of thumbnails from all of the pages in a document, using row-major order, on a single screen. No scrolling is ever required to view all of the thumbnails—they are scaled accordingly when the number of pages increases. A particular page can be viewed by a simple middle click on the appropriate thumbnail. At this time a short (approximately 100ms) animation scales and moves the page image to the center of the screen. The reverse animation (when a page is zooming out) allows the eye to follow the page to its resting location in the thumbnails. Again, this is initiated by middle clicking on the page. The last visited page is made further identifiable by alpha-blending a red shading onto it, increasing its border size to 3px and changing the border colour to red.

The middle mouse button was chosen for modal switches, as this is already used for scrolling activities, such as Microsoft products that allow rate-based scrolling it be initiated by middle clicking. This choice was carefully considered, as the use of all mouse buttons are already overloaded (operating system and application dependant). Commonly, the left mouse button is used for cursor positioning in text and the right mouse button for a context menus. The middle mouse button could easily be replaced with a key combination if it was deemed more convenient.

3.1.1 Thumbnail Dimensions

The size, position and generation of thumbnail images influences the viability of the Space-Filling Thumbnails system. The thumbnails must be positioned systematically, scaled to the largest size possible and be of high enough quality/resolution to allow features to be distinguished between pages.

The prototype system uses pre-generated thumbnail images. Dynamic generation of thumbnail images is feasible, as demonstrated by Microsoft Word, PowerPoint (in Slide Sorter mode) and Adobe Reader. Two key advantages are gained for the evaluation system by using pre-generated thumbnails instead of dynamically resizing pages. The most important is that it allows us to display higher quality images, due to OpenGL rendering algorithms. The second is more consistent performance, especially important in the evaluation situation.

The size and placement of the thumbnails is crucial to optimise the use of screen real estate. As described earlier, we must place the thumbnails so that all are displayed on a single screen and all maintain their aspect ratio. With these restrictions, we are not always going to be able to fill *all* the available space, but we should endeavour to reduce the amount of wasted space.

We now describe the methodology used for calculating the optimal size of the thumbnail images and hence the number rows and number of columns that should be used.

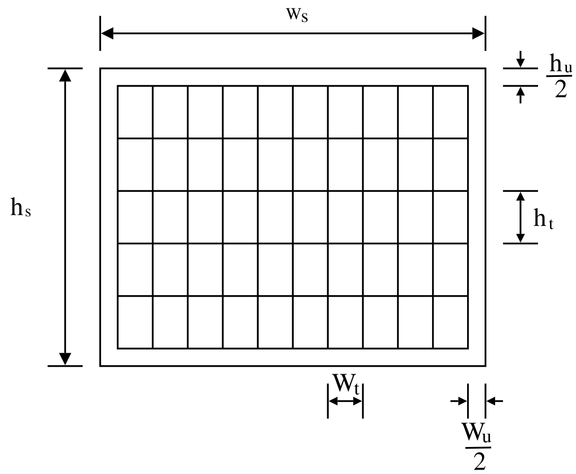


Figure 3.1: Dimensions for thumbnail size calculations

Firstly, we assume that we are always dealing with electronic documents that are produced for paper that is A4 in size. By definition then, a page must maintain a width-to-height aspect ratio of $1 : \sqrt{2}$ [20]. Therefore, our thumbnails must also maintain this ratio.

Figure 3.1 shows diagrammatically the dimensions required. Let us assume we have an available screen size (in pixels) of width w_s and height h_s. To meet the above-mentioned aspect ratio, the complete screen will not always be completely filled. We denote the width of the unused space as w_u and the height of the unused space as h_u. The aim is to reduce both w_u and h_u to the minimum possible so as to increase each of the n_t thumbnail widths, w_t and heights, h_t.

We need to calculate the optimal number of rows n_r and columns n_c, remembering that a thumbnail cannot be split across a row or across a column. All thumbnails must occupy the same area. We select the $max(h_t \times w_t)$ given that:

$$
\begin{aligned}
h_t &= \begin{cases} \frac{h_s}{n_r} & \text{if}(\frac{w_s}{n_c} \times \sqrt{2} > h_s) \\ \sqrt{2} \times w_t & \end{cases} \\
w_t &= \begin{cases} \frac{h_t}{\sqrt{2}} & \text{if}(\frac{w_s}{n_c} \times \sqrt{2} > h_s) \\ \frac{w_s}{n_c} & \end{cases}
\end{aligned}
$$

$$
\text{where } n_c \in \{1 \ldots n_t\},\ n_r = \left\lceil \frac{n_t}{n_c} \right\rceil
$$

The unused space for each dimension can then be calculated as:

$$
\begin{aligned}
h_u &= h_s - (h_t \times n_r) \\
w_u &= w_s - (w_t \times n_c)
\end{aligned}
$$

This unused space is then evenly distributed between either the left and right or top and bottom of the screen. For simplicity these formulae do not take into account a 1px black border that is drawn around each of the thumbnails to provide separation and delineation.

3.1.2 Page Overview

One concern with the SFT system is that the thumbnails will become too small to be easily recognised and distinguished when the document length is large. To ease this problem, a *page overview* tracks the mouse

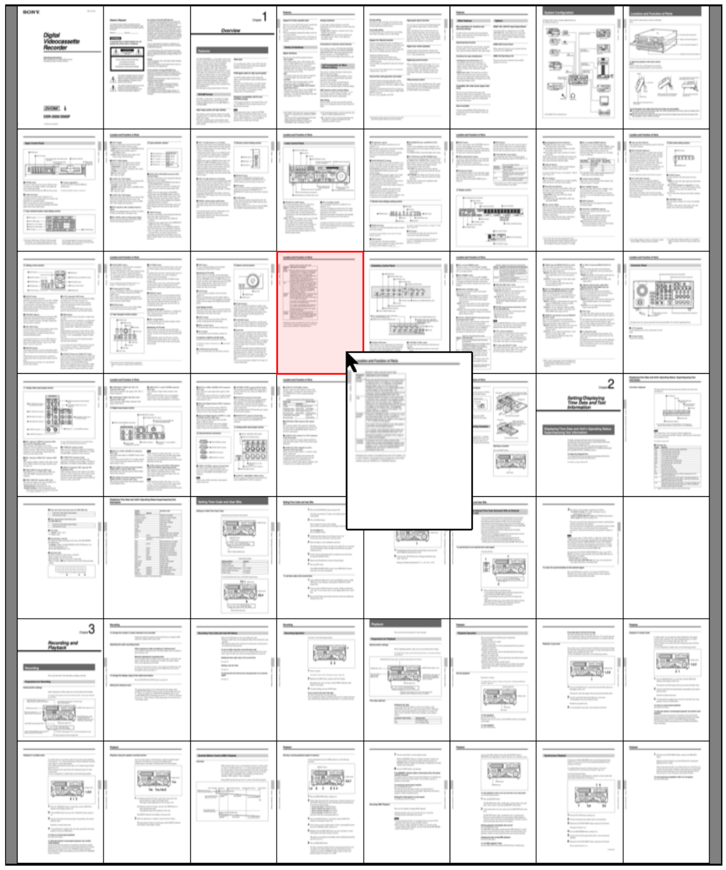


Figure 3.2: The Space-Filling Thumbnails system. The last viewed page is shaded red, with the page overview “hanging” off the cursor.

cursor when the thumbnail height reduces below 200px. The page overview is 154 × 200px in size and its position is adjusted so that it is never offscreen. The SFT system with the page overview is shown in Figure 3.2.

3.2 Competing Interfaces

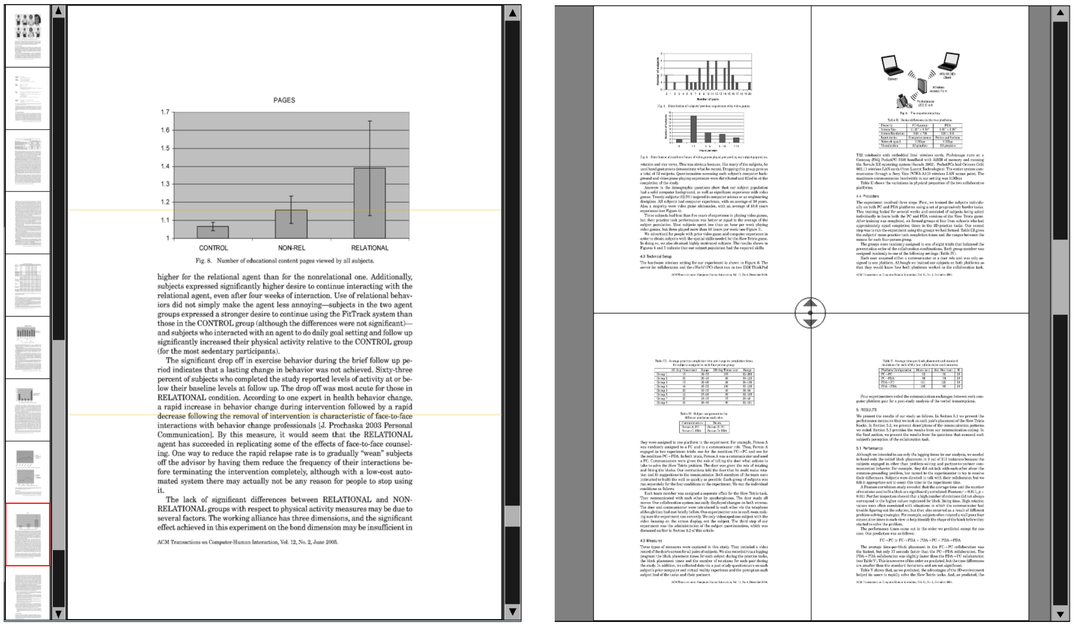
In addition to the Space-Filling Thumbnails system we also implemented six other interfaces for comparative evaluation. Each of these are described here.

3.2.1 Scrollbars

The scrollbar implementation functioned in the same manner as a standard scrollbar system. Users can navigate the document by either manipulating the scroll thumb, scroll trough or the direction arrows.

3.2.2 Thumbnail-Enhanced Scrollbars

The Thumbnail-Enhanced Scrollbars interface includes a column of thumbnails to the left of the main navigation pane. A screenshot of the thumbnail-enhanced scrollbar system is shown in Figure 3.3(a). This implementation operates in the same manner as those on mainstream interfaces, such as Adobe Reader. A user can control the standard scrollbars in the main pane in the normal fashion and the thumbnails will follow this manipulation. The user may also manipulate the scrollbars associated with the thumbnails. The main document does not follow movements in the thumbnails until one is clicked on, at which time the



(a) Thumbnail Enhanced Scrollbars (b) Multi-page Rapid Serial Visual Presentation

Figure 3.3: System Implementation Screenshots

document in the main pane moves to the selected page. The page currently shown on-screen is highlighted in the thumbnails by a 3px red border. The system will always display ten thumbnail images at a time.

3.2.3 Rate-Based Scrolling

The Rate-Based Scrolling system operates in the same manner as that in the Microsoft Windows applications. Scrolling can be performed in one of two manners. Rate-based scrolling can be initiated by middle clicking and then displacing the mouse an appropriate distance to achieve a comfortable scroll rate. Scrolling is halted by a second middle mouse click. Alternatively, one can drag with the middle mouse button again displacing the mouse an appropriate distance. This scrolling is terminated by simply releasing the mouse button.

In this implementation there is a linear relationship between the scroll speed and the displacement of the mouse from the position of initiation. The maximum document scroll rate is 47cm/sec, achieved at a displacement of 170px, as empirically derived by Cockburn and Savage [5].

We provide *passive* scrollbars in this system to allow us to isolate the efficiency of rate-based scrolling. The scrollbars cannot be used for navigation, however it is unrealistic to remove them, as they provide the only spatial cue as to the current position in the document.

3.2.4 Rapid Serial Visual Presentation (RSVP)

Rapid Serial Visual Presentation is a navigation technique that overlays the next page in the document on top of the current page being viewed. Discrete pages are only ever available—one cannot view half of one page and half of the following page. A page is presented for a user controlled amount of time before the next is laid on top. Our RSVP implementation is equivalent to Spence et al's [34] "slide show" mode (see section 2.1.3).

This form of navigation utilises rate-based scrolling to control the speed at which the pages are presented. In this implementation scrolling is initiated by use of the middle mouse button. The distance dragged from the original cursor position determines the *page flipping* speed. A maximum page-flipping speed of 10pages/sec (100ms per page) is attained at a mouse displacement of 170px. This represents the

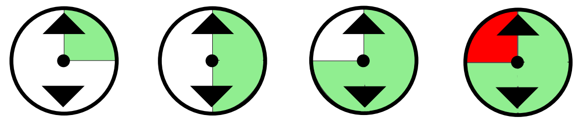


Figure 3.4: Enhanced Rate-Based Scrolling icon for RSVP interfaces, illustrating scrolling toward the end of a document

fastest comfortable rate for cognition, as described by Spence et al. [34].

The system includes an enhanced rate-based scrolling cursor icon to provide the user with an indication as to when the next page will be presented. To demonstrate, a series of screenshots are shown in Figure 3.4, visualising how the enhanced rate-based icon fills over time. The icon fills in a clockwise direction when navigation is directed towards the end of the document, and a counter-clockwise direction when navigation is aimed at the beginning of the document. The rate at which this fills is proportional to the distance the mouse has been dragged from the origin of depression, which is proportional to the rate of page visualisation. The icon is filled with an alpha-blended green colour until reaching the final quarter, at which time it is filled with an alpha-blended red colour to warn users the page is about to be flipped.

The Flipper system implemented by Sun and Guimbretière [35] included a "rewind" feature to counter a user's delay in reacting to seeing the correct page. When using the Flipper interface one can easily "flip" past the required page, as there is no feedback regarding the flipping-rate or when the next page is to be presented. The rewind feature scrolls in the opposite direction to that of the most recent scrolling action, according to the speed at which one was scrolling and accounting for standard reaction times. We believe that our enhanced rate-based scrolling cursor negates the need to add this functionality. Rewind features are hard to tune and require a large amount of learning before users become familiar with them. For these reasons we saw no justification to implement this feature. As with the rate-based scrolling interface, RSVP contains passive scrollbars.

3.2.5 Multi-Page Rapid Serial Visual Presentation (MRSVP)

Multi-Page Rapid Serial Visual Presentation (MRSVP) is similar to RSVP except that it displays four pages simultaneously, as shown in Figure 3.3(b). This is to imitate Spence et al's [34] "mixed-mode" presentation, where four images are displayed at a time.

To view a complete page at reading zoom (and to be able to complete a task), the user must left click on one of the four previews. This enlarges the selected page to full size using the same scaling algorithm as SFT. To return to the four preview images the user may either left click or drag with the middle mouse button to initiate scrolling to the next screen of four pages. The parameters for maximum flip-rate are the same as those in the RSVP interface. The enhanced rate-based icon was also used in this system.

3.2.6 Speed Dependent Automatic Zooming (SDAZ)

Speed Dependent Automatic Zooming is a technique that automatically zooms out from a document as the scrolling speed increases (see Figure 2.2). By scrolling slowly the document stays at normal (reading level) zoom. By scrolling at a fast rate the document will be fully zoomed out, making many pages visible at once.

This form of navigation uses rate-based scrolling to control the speed at which scrolling and therefore zooming occurs. The distance moved from the original cursor position determines the scroll and zoom speed. The two navigation techniques described for rate-based scrolling are available—click and move or press and drag. The document returns to normal zoom when the rate-based scroll is terminated. Again, to isolate the efficiency of the interfaces, the scrollbars are passive. The implementation of this system was a modified version of that created by Cockburn and Savage and hence, the interface is calibrated to the optimal measures they determined in [5].

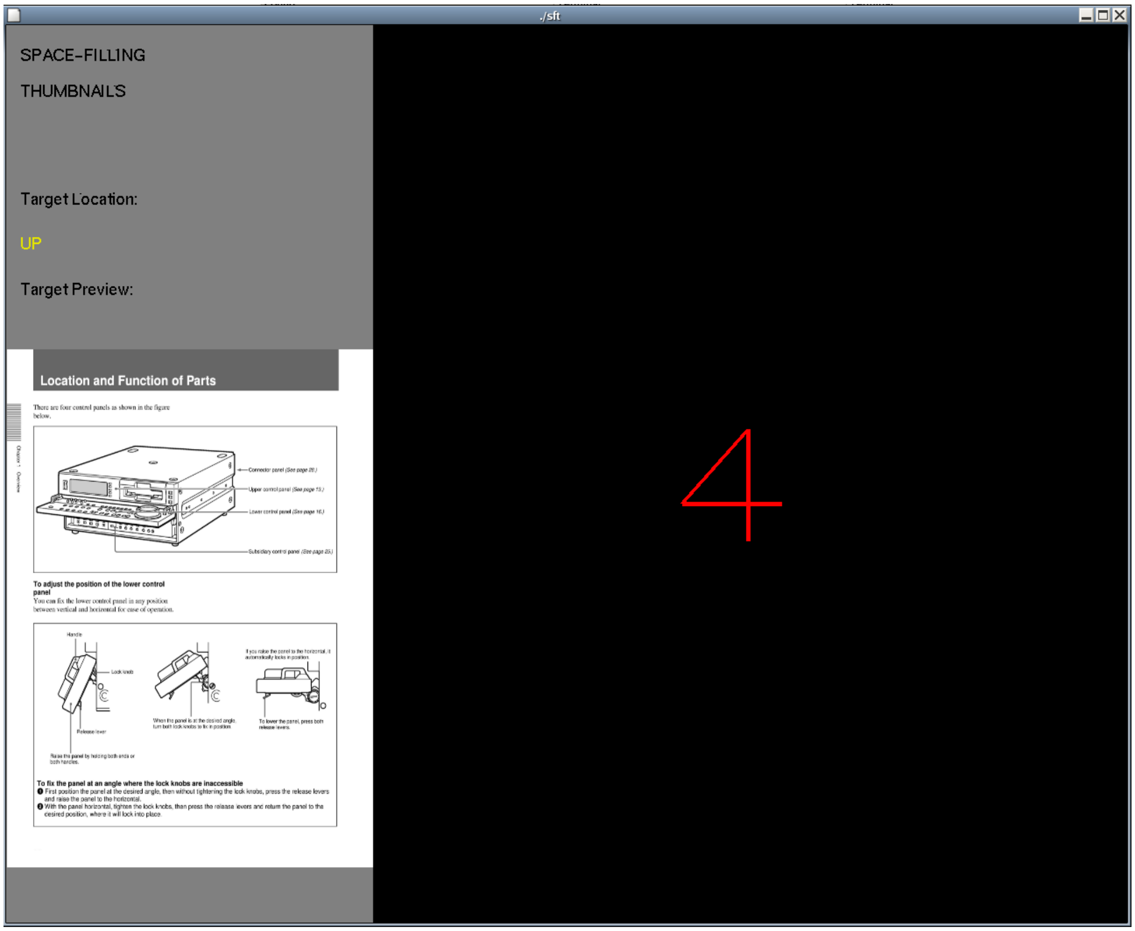


Figure 3.5: Task Cueing Interface

3.3 Experimental Control Software

The evaluation system used for presenting the tasks and interfaces to participants is fully automated. The system prompts the user to begin a task by clicking on an OK button. This provides the user with time to rest beforehand, if required. The user is then given a seven second countdown (as shown in the Figure 3.5) in which they are asked to memorise the target that they will be locating. During the memorisation period, the interface to be used is greyed out, and the mouse pointer disabled. When the search time begins, the interface is automatically revealed and the mouse pointer restored to the center of the screen.

The target page is shown in the cueing area, as seen on the left hand side of Figure 3.5. The cueing area also displays the interface name and the direction of the page *from the starting location*. Note that this direction does not change as the user navigates through the document, otherwise one could use this to locate the correct page in a binary-search manner. The cueing area remains in place throughout the search for the target. The cueing region of the screen has a size of 420×1024px and the interfaces are displayed in remainder.

The user navigates using the interface supplied to locate the required target. To achieve a target, the user is required to have at least one third of the correct page onscreen, and all mouse buttons released. The software automatically detects when the correct page has been reached and the next task introduction screen is presented. The software produces no feedback if the user selects an incorrect page thinking it is the target.

3.4 Experiment Apparatus

All participants in each experiment performed the tasks simultaneously, in a lab of identical specification machines. The experiments were run on Intel Pentium 4 2.8GHz computers, equipped with 1GB RAM and NVIDIA GeForce FX5200 graphics cards connected to 19-inch CRT displays at 1280x1024 resolution and 75Hz. All input was through a Logitech three-button opto-mechanical mouse with a one-to-one control-display gain mapping.

4 Formal Evaluations

It essential in the proposal of any new system design that it is evaluated to ensure that it performs as expected. A series of experiments was conducted to evaluate the efficiency and subjective satisfaction of the Space-Filling Thumbnails document navigation system against that of the six competitor systems.

Experiment One describes a seven interface "scroll-off" in which we compare Space-Filling Thumbnails with the six other implemented interfaces. Experiment Two delves deeper into the robustness of SFT by comparing it with the "best of the rest" interface, Thumbnail-Enhanced Scrollbars, using documents of various lengths. The third experiment investigates the possible weakness of SFT, long documents with few distinguishable features.

4.1 Experiment One — Finding a Challenger

This is the first experiment where the Space-Filling Thumbnails interface is "let-lose" on volunteer participants. Their reactions and comments will be invaluable for improving the interface for the further experiments. In this experiment we will isolate one interface that will be used for further comparative studies with SFT.

4.1.1 Experiment Aim

Experiment One has two main goals. The first is to confirm the viability of the SFT interface. We need to ensure that the system was usable and that it was possible to locate specific pages using it. If the interface does not perform as expected then issues raised will need to be addressed before further evaluations can proceed.

The second goal is to find a "best of the rest" interface for our subsequent experiments. In order to conduct more in-depth latter experiments, we will reduce the number of interfaces to just two, allowing us to evaluate various stereotypical documents and their lengths.

4.1.2 Hypothesis

In this experiment we predict that the Space-Filling Thumbnails interface will have the fastest times for both visual and spatial search over all other interfaces.

The design of the SFT system allows users to view a complete overview of a document in one screen. It allows easy scanning of the pages available, making for an efficient visual search. Recall that the SFT interface exploits a user's spatial memory, so we expect it to perform even better when the spatial location of a page is implanted into a participants memory. Every page in the SFT interface stays in the same position in the 2D matrix, so returning to pages previously visited should be fast.

It is difficult to predict the performance of other interfaces. All participants will be familiar with the Scrollbar, Thumbnail-Enhanced Scrollbar and Rate-based Scrolling interfaces so this may influence their performance. However, SDAZ and RSVP based interfaces have been shown by other researchers to be faster than their mainstream counterparts.

4.1.3 Participants

The thirteen volunteer participants (one female and 12 males) in this experiment were all Postgraduate students in the Computer Science and Software Engineering department at the University of Canterbury.

4.1.4 Experimental Design and Procedure

Interface Familiarisation

A demonstration was given to participants before they begun the experiment. In this demonstration seven target pages were located, one with each of the interfaces. This was intended to show participants both the experimental interfaces and cueing mechanisms. All participants were given practise with each of the interfaces before beginning the series of tasks with that interface. One minute "free practise" was given, with users encouraged to simply view pages in the document and become familiar with the interface controls. They were then given four training tasks, that were identical in their nature to those they would be receiving in the timed evaluation. Participants were provided with an instruction manual and were encouraged to ask if they were unsure of the interface controls before beginning. This training was essential to eliminate the learning effects that can occur when using new systems.

Note that in all of these systems the *only* way to navigate through a document was by using the controls described in Chapter 3. Auxiliary movement techniques that are provided in production systems, such as the `Page Up` and `Page Down` key mappings were not included. This is so that we can isolate the behaviour of each of the systems.

Experimental Procedure

This experiment involved two distinct tasks for each of the target pages. The initial search for the target page was a pure visual search—the only thing known about the location of the page was whether it was up or down from the starting position. The same page was then immediately re-acquired. Participants were informed of this at the beginning of the experiment. They were encouraged to try to remember the location of the page, so that it could be found as quickly as possible the second time around. This tactic was used to artificially implant the spatial location of the page in the participants memory. This simulates the situation where a user is familiar with a document and is frequently visiting a particular page in that document. All tasks were completed with one interface before beginning with the next. At the conclusion of the experiment we asked participants to complete a subjective evaluation form.

Documents

Eight documents were used in total, one of which was used for all of the training tasks. The remaining seven were used for the evaluation tasks, one for each of the interfaces. The order in which participants viewed the documents remained constant, however the order the interfaces were exposed to the participants was balanced using an incomplete Latin square. All of the documents were journal papers selected from the *ACM Transactions on Computer-Human Interaction* (TOCHI), any that were over 30 pages were truncated by removing pages from the end. The page numbers in the documents were removed to prevent participants using these to aid their search for the target page. The documents were also converted to grey-scale to avoid colour cues being used to aid visual search. References to the documents used can be found in Appendix A.

Target Pages

The five target pages for each document were selected at random, ensuring they were within the first or last nine pages. The starting location for each task was also randomly selected to be 15–18 pages from the target page. These constraints were not made known to participants.

Experimental Design

The primary dependant variable is task completion time: the elapsed time from revealing the interface to location of the correct target page. The experiment was run as a 7×2 repeated measures analysis of variance (ANOVA). There are seven levels for the within subjects factor *interface*: MRSVP, Rate-based Scrolling (RBS), RSVP, Scrollbars, SFT, SDAZ and Thumbnail-Enhanced Scrollbars (TES). The within subjects factor *search iteration* has two levels: visual and spatial search.

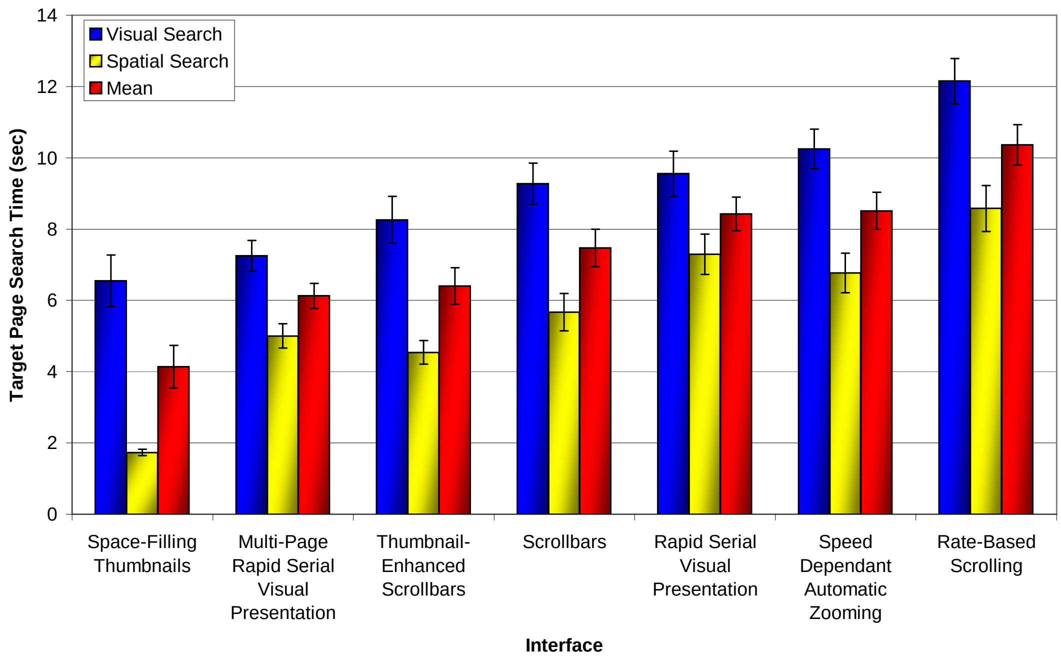


Figure 4.1: Experiment One: Target page search times

4.1.5 Results

We divide the results from this experiment in two—empirical and subjective results—these are summarised in the following sections.

Empirical Results

A summary of the empirical results is shown in Figure 4.1. Each interface has the visual search, spatial-find and mean times shown. Search times for this experiment were capped at thirty seconds, resulting in the removal of six of the 910 tasks. There is a significant main effect for the factor interface ($F_{6,72} = 24.65$, $p < 0.01$). The Space-Filling Thumbnails interface allows the fastest mean search time of 4.1secs (standard deviation 3.0), followed by MRSVP (6.1secs, s.d. 1.8) and Thumbnail-Enhanced Scrollbars (6.4secs, s.d 2.6). A post-hoc Tukey test gives an Honest Significant Difference (HSD) of 2.47secs ($\alpha = 0.05$).

The analysis of the interaction between the search iteration and interface also yields a significant main effect ($F_{6,62} = 2.6$, $p < 0.05$). Space-Filling Thumbnails has the fastest visual search time of 6.5secs (s.d. 2.6), followed by MRSVP with 7.3 secs (s.d. 1.5) and Thumbnail-Enhanced Scrollbars (8.3secs, s.d. 2.4).

In the spatial-recall task the SFT interface out-performs all others. SFT allows the fastest mean spatial-find, with a time of 1.7secs (s.d. 0.3), with the runner-up being Thumbnail-Enhanced Scrollbars (4.5secs, s.d. 1.2), followed by MRSVP (5.0secs, s.d. 1.2). Figure 4.2 compares the interfaces with regard to their performance increase for the spatial search, as a percentage of the visual search time for that interface. The decrease in spatial search time over the original visual search is more than 70% for SFT. We see that SFT aids spatial re-acquisition 30% more than its nearest competitor TES.

Subjective Results

Participants were asked to uniquely rank the interfaces in the order that "best helped you to satisfy the required tasks" (ranks 1-7, best-worst). Space-Filling Thumbnails had the best ranking, with a median of 1st and Thumbnail-Enhanced Scrollbars was second. RSVP and MRSVP had an equal last ranking of 6th. The complete results are shown in Table 4.1. The poor rating of the RSVP based interfaces were backed

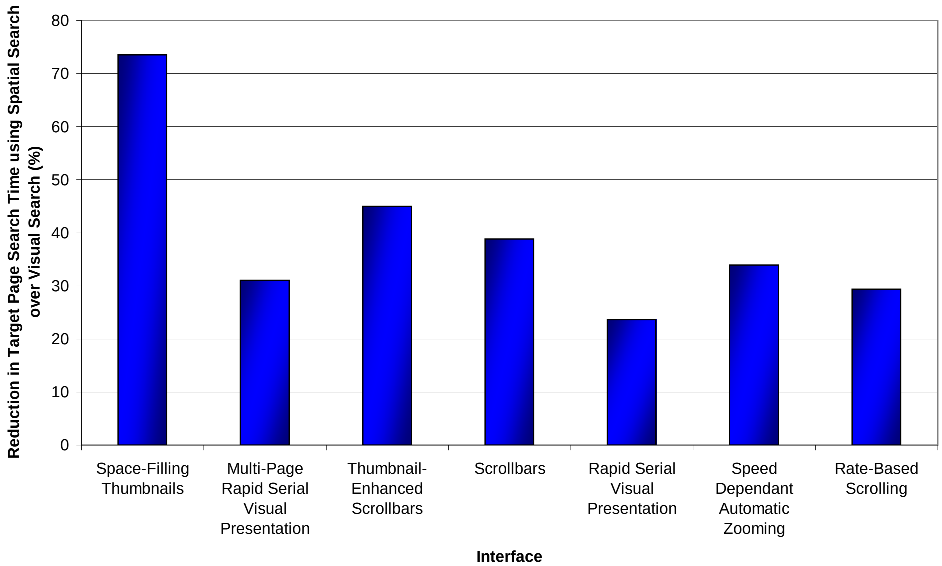


Figure 4.2: Experiment One: Reduction in search time for spatial search

	Preference Ranking (cumulative %)							Search Efficiency* (mean (s.d))	
	1st	2nd	3rd	4th	5th	6th	7th	Visual	Spatial
MRSVP	0 (0)	1 (8)	0 (8)	2 (23)	2 (39)	4 (69)	4 (100)	2.86 (0.99)	3.07 (0.80)
RSVP	0 (0)	1 (8)	1 (15)	1 (23)	0 (23)	5 (62)	5 (100)	2.86 (1.12)	3.21 (0.86)
RBS	0 (0)	0 (0)	3 (23)	1 (30)	5 (69)	2 (85)	2 (100)	3.07 (0.96)	3.64 (0.48)
Scrollbars	0 (0)	2 (15)	4 (46)	4 (77)	1 (85)	1 (92)	1 (100)	3.36 (0.97)	3.86 (0.74)
SFT	10 (77)	0 (77)	1 (85)	1 (92)	1 (100)	0 (100)	0 (100)	4.07 (0.88)	4.86 (0.35)
SDAZ	2 (15)	2 (30)	2 (46)	3 (69)	3 (92)	0 (92)	1 (100)	3.50 (0.82)	3.71 (1.03)
TES	1 (8)	7 (62)	2 (77)	1 (85)	1 (92)	1 (100)	0 (100)	3.79 (0.56)	4.07 (0.80)

Table 4.1: Experiment One: Subjective rankings for preference and efficiency of interfaces. *0 Disagree, 5 Agree

up with negative comments such as “it felt like I was about to get a headache, or a seizure”. Comments regarding SFT were positive—“very easy to remember rough location of thumbnail,” although some were speculative about its performance in more extreme conditions: “more pages means thumbs need to be smaller, which means harder to find pages”.

Significant subjective differences between the interfaces resulted when participants were asked whether the “interface was efficient for finding” either the first page (Friedman $\chi^2 = 17.9$, $p < 0.01$) or second page ($\chi^2 = 30.2$, $p < 0.01$) of the pair. For the visual search task, SFT had a mean of 4.15 (s.d. 0.9) followed by Thumbnail-Enhanced Scrollbars with a mean of 3.77 (s.d 0.6). For the spatial search task SFT had mean of 4.9 (s.d. 0.3) with Thumbnail-Enhanced Scrollbars having a mean of 4.0 (s.d. 0.8).

4.1.6 Discussion

We present here a discussion of the results obtained in this experiment and then look at comparing these results with those reported by others and finally investigate page-flipping intervals for RSVP-based interfaces.

Figure 4.1 showed that the Space-Filling Thumbnails interface had the fastest mean search time in both iteration conditions, validating our hypothesis. Our prediction that spatial search in SFT will be fast was also demonstrated to be correct. Second fastest was MRSVP, followed by TES whose overall means are not separated by the HSD of 2.47secs described above. TES produced a slower visual search time, but in the subsequent spatial search out-performs MRSVP.

The subjective results for these two interfaces influence the choice of challenger for the second experiment. The subjective results showed that TES was ranked second for "best satisfying the required tasks", while MRSVP, along with RSVP were ranked 6th (last) equal. A common remark made by participants was that the rapid changing of pages required intense concentration and that it was difficult to pinpoint the correct page.

Taking these results into account, we chose Thumbnail-Enhanced Scrollbars as the competitor interface for Experiment Two. It is commonly acknowledged by researchers that interfaces that people do not like, they will not use. Further, to produce credible and *comparable* results we should compare SFT with a document navigation interface that is common in the "real-world". It is for these reasons that the Thumbnail-Enhanced Scrollbars interface is chosen as the interface for comparison in the forthcoming experiments.

Comparison of Results

Sun and Guimbretière [35] compared their Flipper interface with SDAZ and "scrolling with thumbnails" which is equivalent to our Thumbnail-Enhanced Scrollbars condition. Our visual search task is only comparable to their "Image Search" task. We are unable to find a significant difference between SDAZ and TES, agreeing with their result. Interestingly the results from the RSVP interface that is similar to their Flipper interface (without a small portion of zooming) also show no significant difference, disagreeing with their results. There are several reasons for this disparity. Firstly, their search tasks were purely for images, whereas our tasks involved pages that contained either text or images or both, emulating a more realistic visual search. Secondly, our RSVP interface and their Flipper interface differ in operation slightly.

Cockburn and Savage [6, 5] found that SDAZ was significantly faster than scrollbars. Again only considering our visual search results, we cannot determine any significant difference, therefore we cannot confirm or rebut their results. Their tasks also involved searching for large distinguishable items—headings or images, as opposed to some of our searches being for pages containing only text.

Hornbæk and Frøkjær's comparison of linear (scrollbars), fisheye and overview+detail (thumbnail-enhanced scrollbars) [16], showed that subjects prefer the overview+detail interface, matching our results. We observed Thumbnail-Enhanced Scrollbars being ranked second only to SFT. They also found that the overview+detail interface slowed participants down. This difference can be put down to the type of task given, with our participants knowing exactly what they were searching for, and their participants performing an essay writing task, which required a more in-depth knowledge of the document.

Characterisation of Rapid Serial Visual Presentation Techniques

We look now at characterising the page flipping rate of both of RSVP based interfaces.

Although much studied for image presentation, few studies have looked at RSVP for document navigation. The maximum recommended rate for page flipping, as suggested by Spence et al. [34] is 100ms. We allowed participants to control the rate at which pages were flipped up to this maximum. Figure 4.3 shows the average flip intervals participants used in their navigation comparing both RSVP and MRSVP. Interestingly, all of the means are over three times this maximal rate of 100ms suggested in [34]. This result is explainable by the significantly larger amount of data displayed on one page of a document compared to an image. Also the likeness in structure of pages required participants to carefully check them.

One may have expected the MRSVP flip-intervals to be four times larger than the RSVP flip-intervals, due to there being four times the information. This turned out not to be the case. In their experiments Spence et al. [34] displayed mixed-mode RSVP (equivalent to our MRSVP) screens for four times longer than their slide-show mode (our RSVP). For this reason one may have assumed that participants would have a flip-interval fours times larger in the MRSVP interface over the RSVP interface. Clearly Figure 4.3 shows this is not the case. One can conjecture that the flip intervals for RSVP were not the "fastest possible", but

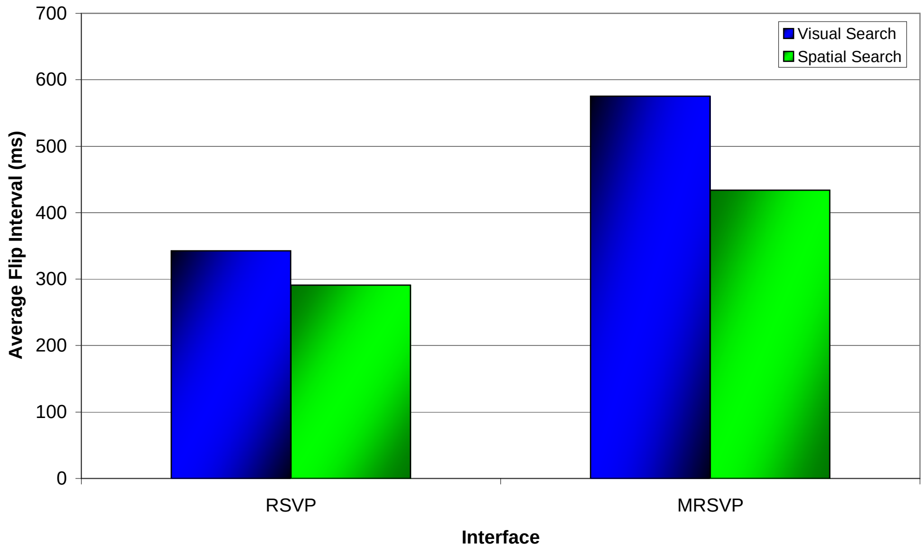


Figure 4.3: Rapid Serial Visual Presentation Techniques: Average flip intervals

rather intervals at which participants were most comfortable. In their eye-gaze analysis of their mixed-mode RSVP (recall, equivalent to our MRSVP), Spence et al. [34] noted that eye movements were focused around the centre of the quadrant, rather than scanning from image-to-image. This may offer some further explanation as to why the MRSVP flip intervals are not as large as expected.

4.2 Experiment Two — Comparing SFT with TES

In Experiment One, Space-Filling Thumbnails followed our predictions for both the visual and spatial searches for the target pages. We concluded that the Thumbnail-Enhanced Scrollbar interface was the "best of the rest" and should be used for further comparative studies. In this experiment we provide a more thorough comparison between the SFT interface and the TES interface. We are interested in three factors in this experiment: the interface type, the document length and how the interfaces perform when spatial location information is not artificiality implanted into the participants memory.

4.2.1 Experimental Aim

In this experiment we aim to determine how each of the two interfaces aid a user's spatial memory in re-acquiring a recently visited page. We wish to investigate how the length of the document effects the search time, to ensure that SFT does not suffer unacceptable performance penalties using long documents.

4.2.2 Hypothesis

Based on the results from Experiment One, we hypothesise that the SFT interface will be faster for all iterations of search. The visual search iteration, as we have seen in Experiment One is significantly faster than other interfaces. The spatial search tasks should follow suit, as SFT allows the learning of page location faster than the TES interface (due to its exploitation of spatial memory). The increasing document length will see an increase in search time, especially in the linear TES interface.

4.2.3 Participants

The participants in this experiment were all volunteers from the Experimental Computer Science (COSC209) undergraduate paper instructed by the Department of Computer Science and Software Engineering at the

University of Canterbury. There were two female and 30 male participants, with an age range of 18 to 36 (mean 22).

4.2.4 Experimental Design and Procedure

This experiment involved participants locating target pages in documents of differing lengths. Participants were asked to locate the same page multiple non-consecutive times so as to determine how each of the interfaces utilises spatial memory to allow quick returning to pages previously visited.

Documents

Three stereotypical types of documents were chosen to give us the three levels of the length factor. The first were conference papers from the *Conference on Computer-Human Interaction* (CHI). These papers are typically 10 pages in length. The second type were journal articles from the *ACM Transactions on Computer-Human Interaction* (TOCHI). The journal articles are typically 30 pages in length. The third document type was an instruction manual, which are variable lengths, but typically around 150 pages. Each of the document lengths were truncated (if necessary) to the values described here. As per experiment one references for the documents used can be found in Appendix A. Page numbers, contents pages and indexes were removed (if applicable) from the documents to prevent participants from using these to aid their search for the required pages. All pages were also transformed to grey-scale to prevent target page search from being aided by colour.

Target Pages

The target pages were randomly selected within the middle third of the document. The starting pages were randomly generated to between 23–33% of the document length either above or below the target page. Participants were unaware of these constraints.

Page Search

Unlike Experiment One, we did not inform the participants that they would be locating the same page multiple times. We also did not ask them to find the same page consecutively. All pages were found for the first time, the order balanced and then all were found a second time. The experiment was run in this fashion so that we could measure the difference in learning of the location of pages between the two interfaces. The first page search will be purely a visual search (the only clue to its location being whether the page is up or down from the starting position). It is envisioned that the second page search will be partially visual search and partially spatial memory retrieval. It is hoped by the third page search it will be a pure spatial memory retrieval.

Participants completed all tasks with one interface before moving to the other. The order in which the interfaces were presented was balanced between participants.

Interface Familiarisation

Before beginning the tasks with each interface a training session was provided. As in Experiment One both of the interfaces were demonstrated and a free-practice session provided. Four specific training tasks were provided to ensure participants were familiar with the cueing mechanisms. A short explanatory manual was also provided, with participants encouraged to ask if they were unsure of the operation of either of the interfaces.

Experimental Design

The primary dependant variable is task completion time. The experiment was run as a $2 \times 3 \times 3$ repeated measures analysis of variance (ANOVA). The within-subjects factor *interface* has two levels: Space-Filling Thumbnails and Thumbnail-Enhanced Scrollbars. The within-subjects factor *document length* has three levels: 10, 30 and 150 pages. Within-subjects factor *iteration* has three levels: first, second and third target page search.

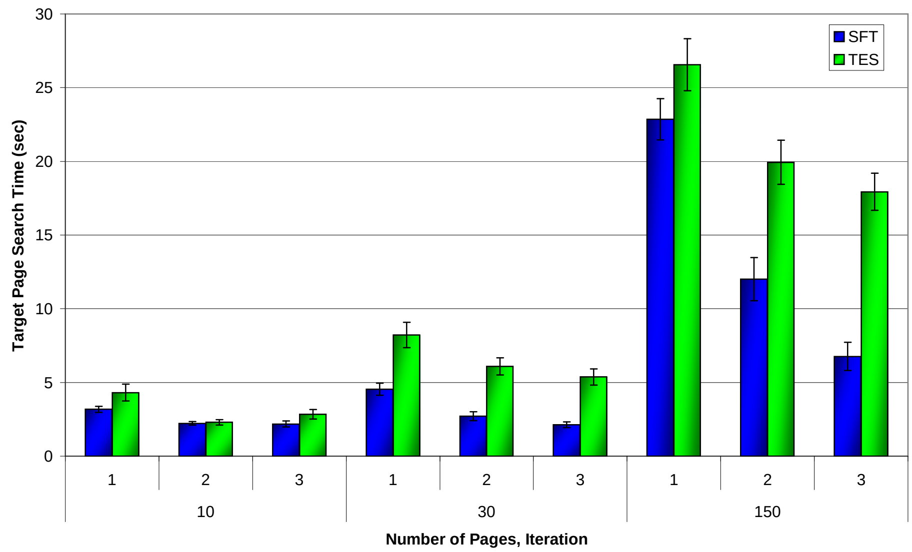


Figure 4.4: Experiment Two: Target page location time by document length

Subjective Evaluation

At the completion of all tasks with each of the interfaces, participants were requested to fill out a subjective measures assessment form using the NASA-TLX [14] subjective workload evaluation. At the conclusion of the experiment participants were requested to fill in a further subjective measures form that compared the two interfaces.

4.2.5 Results

Empirical Results

A summary of the empirical results from this experiment can be found in Figure 4.4. Tasks for this experiment were capped at 60 seconds, resulting 25 tasks for SFT and 26 tasks for TES being discarded from the original 1728 recorded times. The SFT interface was significantly faster than the TES interface with overall mean times of 6.5secs (s.d 7.8) and 10.4secs (s.d 10.0) respectively ($F_{1,31} = 66.6$, $p < 0.01$). The large standard deviations are explainable due to the large variance in times taken to locate the target pages in the 150 page document.

There is a significant main effect for the document length factor ($F_{2,62} = 348.67$, $p < 0.01$). A post-hoc Tukey test gives an HSD of 3.6 secs ($\alpha = 0.05$), indicating that the 150 page document takes significantly longer than the 10 and 30 page documents. As Figure 4.4 shows the SFT search times for the 10 and 30 page documents differ only slightly.

Subjective Results

The results of the NASA-TLX subjective workload assessment questionnaires are summarised in Figure 4.5. The only significant difference is for the *Physical Demand* category, where SFT has a mean of 2.31 (s.d. 1.13) and TES has a mean of 2.82 (s.d 1.31) (Paired T-Test, $T_{33} = 3.59$, $p < 0.05$).

The participants answered subjective questions regarding their performance with each of the interfaces. When asked to rate how efficient each of the interfaces were for the tasks on a 5-point Likert scale, a significant difference resulted, with SFT having a mean of 3.85 (s.d. 1.03) and TES having a mean of 3.37 (s.d. 0.94), (Wilcoxon $z = 1.76$, $p < 0.05$). In response to the Likert scale question "I liked using", SFT had a mean rating of 3.9 (s.d. 1.1) and TES received a mean rating of 3.4 (s.d. 0.9) indicating a significant

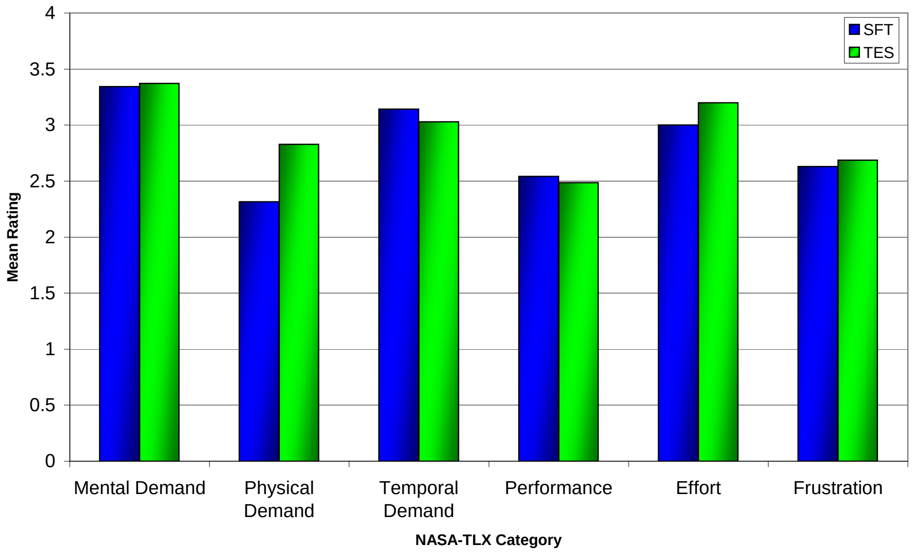


Figure 4.5: Experiment Two: NASA-TLX subjective workload assessment results

difference (Wilcoxon $z = 1.88$, $p < 0.05$).

There was no reliable difference in responses to the 5-point Likert scale question "It was easy to return to previously visited pages," meaning we cannot judge whether participants recognised that they were aware of the decrease in time spent searching for a page that they had already seen using Space-Filling Thumbnails. As we expected participants recognised that the Space-Filling Thumbnails interfaces made it easier to remember the location of pages. When asked to indicate on a Likert scale whether "I could remember the location of pages," Space-Filling Thumbnails had a mean of 3.9 (s.d. 1.3) and Thumbnail-Enhanced Scrollbars had a mean of 3.2 (s.d. 1.1), showing a significant difference (Wilcoxon $z = 2.6$, $p < 0.01$).

Finally we asked participants to indicate overall which of the interfaces they preferred. Of all the participants 70% indicated that they preferred the Space-Filling Thumbnails interface, giving a significant result ($\chi^2 = 3.8$, $p < 0.05$).

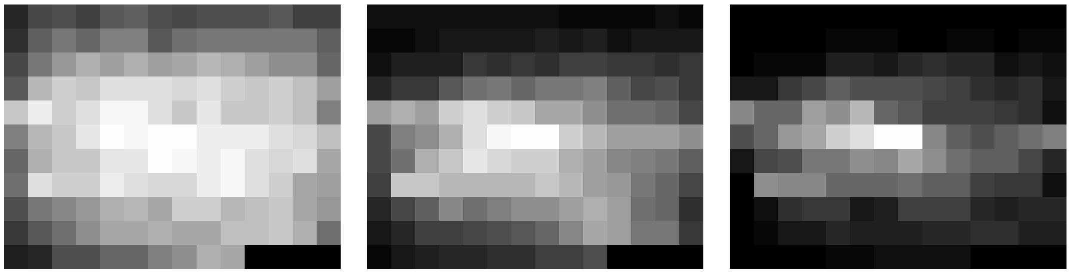
4.2.6 Discussion

We present a brief discussion of these results, followed by an analysis of the use of the page overview in the SFT tasks in the 150 page documents.

From this experiment we see that the SFT system is robust and is faster than its competitor TES in all of the tested document lengths. Interestingly, for both the 10 and 30 page documents, the search time in the third iteration using Space-Filling Thumbnails are identical. This is tribute to how easily pages can be located using this interface. This indicates that when a user is familiar with the document, the search time is not proportional to the length of the document. This is due to the "random-access" nature of Space-Filling Thumbnails, as opposed to the linear navigation required in Thumbnail-Enhanced Scrollbars (and most other previously proposed interfaces).

From Figure 4.4 we see that the biggest gain in spatial memory occurs in the large 150 page document. The final search for the target pages takes 30% of the time of the original visual search. With the TES interface the final search takes 67.5% of the time of the original search. This can be put down to two factors: Firstly, the effort expended initially searching for a page is greater in the larger document. A user is more likely to utilise their spatial memory in this situation than if the search is easy. Secondly, with the smaller documents, it is likely that the search time is approaching that predicted by Fitts' Law [11], limiting the minimum acquisition time (this is discussed further in Chapter 5).

This experiment shows that while increasing the document length *does* increase the amount of time



(a) Iteration One (b) Iteration Two (c) Iteration Three

Figure 4.6: Use of Page Overview During Search (white = heavy use)

required to search for a particular page, the SFT interface still performs better than the TES interface with all document lengths. With the reduction in thumbnail size as the document length increases there was the concern that the performance with SFT may degrade unacceptably. To confirm that even with book length documents SFT is viable, experiment three will investigate the effect of searching in a long, landmark-less document. A discussion on the Hick-Hyman Law for Choice reaction for the Space-Filling Thumbnails interface as the number of pages increases is presented in Section 5.3.5.

The significant difference noted in the physical demand category of the NASA-TLX workload assessments is expected. The SFT interface requires far less physical movement, clicking and dragging than the Thumbnail-Enhanced Scrollbars interface requires to obtain the target page. The act of dragging a scrollbar is more strenuous than simply moving the mouse.

Characterisation of SFT Interaction

It is interesting to look at how the page overview of the SFT interface is being used to aid the visual search of the document. The page overview is not necessary in the 10 and 30 page documents as the thumbnails are large enough. However, it is used in the 150 page document. Figure 4.6 shows a visual representation of the use of the page overview tool. Each square in the figure represents one page in the display of 150 thumbnails. A white colour indicates that all 32 of the participants viewed that page with the overview at some point, and black colour that no-one viewed that page, with the shades of grey representing some proportion in-between. We see in the first iteration, (Figure 4.6(a)) that majority of people viewed majority of the pages in the document using the page overview. In iterations two and three (Figures 4.6(b) and 4.6(c) respectively) the mouse movements (and hence page overviews) are much more confined and directed towards the target pages, indicating participants are performing less searching and relying more on their memory, as was seen in our empirical results. Note that the cursor originally starts in the center of the screen, accounting for some of the concentration in the center of the figures.

4.3 Experiment Three — Space-Filling Thumbnails' Worse Case Scenario

Several of the participants in the first experiment expressed their concern that when the number of pages became large, the size of the thumbnails would make them indistinguishable. Comments were made such as "more pages means thumbs need to be smaller, which means harder to find pages" and "good when not too many pages—could get rapidly worse with more". In this third and final experiment we investigate the performance of Space-Filling Thumbnails when it is used with a large document (300 pages). We are also interested in how the Space-Filling Thumbnails technique performs when pages in the document are moved—how the search time for them degrades and whether it degrades (in the worst case) to a visual search task.

4.3.1 Experimental Aim

The aim of the third experiment is to allay some of the fears that the system may be unusable when a document with a large number of pages is viewed. We also wish to determine how moving pages (to imitate editing) effects re-acquisition time.

4.3.2 Hypothesis

In this experiment we expect the target pages to be a lot more difficult to find than in previous experiments, due to the length and lack of distinct features in the document. We expect visual search time to still be faster in SFT than in TES as the general structure of the pages will still be able to be made out. Also, TES will involve large amounts of scrolling to achieve the correct targets. As per previous experiments the spatial search in SFT should be significantly faster than the visual search of SFT and the spatial search in TES.

When pages are moved in the document we expect to see an increase in search time for both interfaces. When using SFT, participants should recall the general area of the page, allowing a fast visual scan of that area, and hence quick retrieval. We expect search times in the TES interface to be approximately the same as the spatial search, as pages will have moved a small distance linearly—a minor change in location in the context of the 300 page document.

4.3.3 Participants

This experiment was run in the same session as Experiment Two, those who were willing were asked to perform these tasks. These participants had immediately prior been using the TES and SFT systems so no training was required. The tasks in this experiment were completed by 12 participants.

4.3.4 Experimental Design and Procedure

There were two distinct tasks sets to be completed with each interface. Firstly, we asked participants to locate pages within a large document with no distinguishable features. Secondly, using that same document, we nudged the same target pages to imitate document editing.

Document

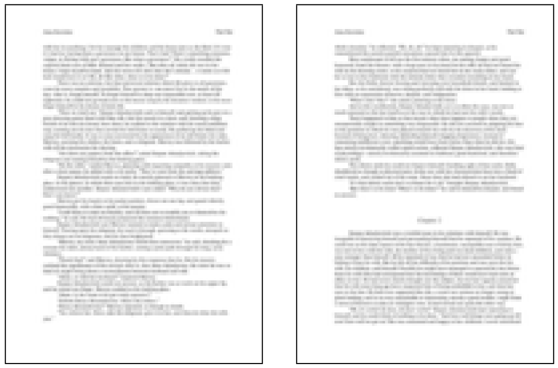
We used the book *Anna Karenina* (see Appendix A), which was truncated to 300 pages. It lacks large distinct objects such as images, graphs and tables. The most distinguishable features in this book were the chapter breaks within the text. Chapters headings were highlighted with colour, so the document was converted to grey-scale and page numbers removed. To demonstrate, two thumbnails are shown in Figure 4.7. One showing a typical page of text (Figure 4.7(a)), and one showing a page that contains a chapter break (Figure 4.7(b)). Note that text cannot be made out, but page structure can easily be seen. We asked participants to locate (and later relocate) pages of both these types.

Targets

In previous experiments the system would automatically move to the next task when at least one third of the correct page was on screen (and all mouse buttons were released). To ensure that participants did not rely on the system to do this identification for them, a one second delay was added before the system would recognise that the correct page had been selected. This prevented participants from using "rapid-fire" selection of pages that were possible candidates.

The same target selection rules were applied as in Experiment Two. Target pages were randomly selected to be in the middle third of the book, with the starting locations randomly selected to be 23–33% of the books length above or below the target page. Half of the target pages were selected as *distinct* pages and half were chosen as *non-distinct* pages.

For the second set of tasks participants were told the target pages had moved up to five pages from their original positions (all targets were actually moved four pages, but participants were unaware of this). All



(a) Pure text thumbnail (b) Thumbnail with chapter break

Figure 4.7: Example thumbnails from Anna Karenina

of the pages in the document were moved either towards the beginning, or towards the end. Excess pages were wrapped to the end and beginning respectively. None of the wrapped pages were target pages.

Participants completed all of tasks using one interface before moving to the next. This included the tasks in which pages had been moved. The order the interfaces were presented was balanced between participants.

Experimental Design

As per previous experiments, the primary dependant variable is task completion time: the elapsed time taken from exposure to the interface to the location of the correct target page. For the first set of tasks, the experiment was run as a $2\times2\times2$ within subjects analysis of variance (ANOVA). The within subjects factor *page type* has two levels: Distinct or Non-Distinct. The within subjects factor *interface* has two levels: Space-Filling Thumbnails and Thumbnail-Enhanced Scrollbars. The within subjects factor *iteration* has two levels: visual search and spatial search. In this experiment, as with Experiment Two, the participants were *not* told that they would be finding the same page again (however, one can safely assume that the participants realised this may happen).

For the second set of tasks where the pages have been moved, the primary dependant variable was again the task completion time. These tasks were run as a 2×2 analysis of variance (ANOVA). The within subjects factors page type and interface are as described for task set one.

4.3.5 Results

The results for both task sets are summarised in Figure 4.8. The large error bars on the figure are a result of the large variance in time taken to retrieve the required pages. The data was log transformed to reduce this variance and allow conclusions to be drawn.

Task Set One

Task set one required participants to locate the same pages in the book multiple times. There was a significant difference between the two interfaces for retrieving these pages, with SFT having a mean of 36.0secs (s.d. 45.0) and TES 54.8secs (s.d. 66.1) ($F_{1,11} = 5.0$, $p < 0.05$). This result is as expected, SFT having proved to be faster than TES in all previous experiments.

The page type factor is significant, with the non-distinct pages having a mean search time of 68.7secs (s.d. 71.0) with the distinct pages taking less than a third of this time (22.2secs, s.d. 20.1). There is no significant interaction between the page type and interface factors. This indicates that it is easier to search for pages with a more distinct structure in either of the interfaces, which would be expected.

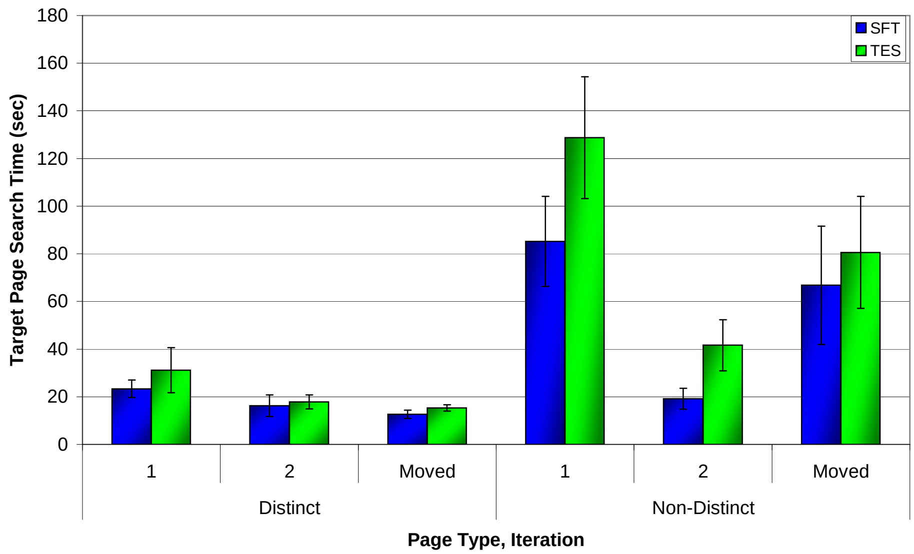


Figure 4.8: Experiment Three: Target page search time using a 300 page book

The iteration factor proved to be significant, with a mean of 67.1secs (s.d. 70.7) for the visual search and a mean of 23.7secs (s.d. 23.8) for the spatial search ($F_{1,11} = 35.68$, $p < 0.01$). The is no reliable interaction between the iteration and interface factors.

Task Set Two

Task set two required participants to locate the same pages as in task set one, except now these had moved. The results from these tasks are also shown in Figure 4.8, displayed as the *moved* iteration. Again the recorded results were extremely variable, so only the *type* factor was significant, with the mean time for non-distinct pages being 73.7secs (s.d. 82.2) and the distinct pages taking less than 20% of that time with a mean of 14.0secs (s.d. 5.4), ($F_{1,11} = 15.8$, $p < 0.01$).

4.3.6 Discussion

We see from these results that SFT is faster in both cases (visual and spatial search), even using this large book. This confirms the first of our hypotheses, that SFT will outperform TES under static page conditions.

In the second part of our hypothesis we predicted that the target location time using TES, for the moved pages condition, would not be largely different from the spatial search time. This is due to all of the document pages being moved linearly. This holds for the distinct page type, but we see *both* interfaces struggle with the moved, non-distinct pages. We can only speculate as to why this is—there is only subtle differences in page structure between the non-distinct pages, obviously making it hard to pinpoint the exact page being searched for. Another possibility is that us informing the participants that the pages had moved may have prompted them to 'throw away' or ignore their spatial memory from previous searches and effectively conduct another spatial search. A point to note is that none of the target pages changed rows in the SFT interface, but one could envision that this would seriously upset a user's spatial memory. We believe that visual search of SFT does not following in the same manner as reading a book (left to right, return, left to right), but rather in a way that reduces eye movement—left to right and then right to left on the next row.

The performance of participants in this experiment varied a large amount (as expected). Documents with fewer pages have larger thumbnails, the features are more distinct and simply recalling the "general" thumbnail position is enough to easily locate the correct target page. However, in these 300 page tasks the lack of distinct features in the pages (recalling that the chapter sections are not as distinct as an object such as a graph, map or table that were in all of the other documents), made the searching tough and to a certain extent some luck may be involved if the participants chose a good strategy.

It is likely that people who have a strong spatial memory performed well in these tasks, and those who did not may have struggled. It would be interesting to be able to grade the strength of participants spatial memory and see if that effects their performance with the interfaces.

Once a document reaches this kind of length, one should be seriously considering whether they should be trying to navigate through it at all. Documents of this length generally contain contents, indexes, and often bookmarks which would likely reduce the need for a complete document search.

5 Discussion and Future Work

In this chapter we discuss possible issues with our evaluations, implications for commercial implementations and possible future work.

5.1 Experimental Concerns

With any experiment there are always concerns as to the validity of the results produced. In this section we discuss some issues relating to our system and experimental method.

5.1.1 Task Cueing Mechanism

As described in section 3.3 tasks were cued by an image of the target page to the left hand side of the screen. This meant that participants could match the search page with the target page. While this is an artificial cue type, we do not believe it favours the Space-Filling Thumbnails interface in any way. This cue is also used to give the most fair comparison between interfaces. The other option would be to use a text-based cue, asking participants to locate a a particular heading, section or larger object such as a graph or map. Often these types of cues can lead to ambiguity, if exact matches are not specified, or confusion if similar headings appear twice.

5.1.2 Page Identification

The system automatically recognised when a participant had at least one third of the target page on-screen and all mouse buttons were released (recall we added a one second delay to this in experiment three). This does not require the participant to explicitly inform the system when they believe they have found the correct page. This may have lead to a small amount of the pages being found by luck. We believe that we obtained more accurate search times using this strategy than if participants had to "submit" their page. Submission of pages raises several issues, such as, what to do when an incorrect page is submitted. Also it is artificial for people to search for a page and then have indicate they have found it.

5.2 Implications for Production Systems

The prototype system implemented for evaluation purposes was specifically designed so that we could isolate the effects that the Space-Filling Thumbnails technique would have on page search time. For this reason several suggested features that would be practical in a production system were not implemented. These are briefly described here.

5.2.1 Pagination

The ability to page through a document at reading zoom would be desirable (without having to zoom out then back in again). This type of activity is especially applicable when reading a document in a linear fashion. It can be provided by having small navigation buttons, such as those provided by Adobe Reader. The standard key bindings for this operation can be applied, using the `Page Up` and `Page Down` keys. Prolonged depression of either of these interaction techniques would result in a "rapid serial visual presentation".

5.2.2 Scrollbars at Reading Zoom

To reduce the paradigm shift required by users new to Space-Filling Thumbnails, it would be wise to maintain scrollbar widgets at reading zoom. This would give users the ability move small distances in a familiar manner. SFT would initially be added as an optional document navigation technique, much as the *Dynamic Zoom* feature is applied in Adobe Reader.

5.2.3 Page Overview

The enlarged page overview provided when the thumbnails are less than 154×200px in size was added to allow users to view the desired thumbnail at a larger scale for inspection before moving to reading mode. Reaction to this feature was mixed. Participants commented that "sometimes the [page overview] got in the road of scanning the thumbnails" and one suggested that the page overview "should have stayed in one position, because it constantly covered up any thumbnails I was trying to look at". For a production system it is likely that this would be an option that could be turned on or off and that the threshold for the size at which it appears would be adjustable.

5.2.4 Resizing the SFT Window

In all of the experiments presented, the navigation interfaces ran in full-screen mode. One issue that arises with SFT is when the window size is altered. The net result of this is that the thumbnails will be moved and resized to be maintain the two key properties: they are space-filling and they fit on one screen. This in turn will result in disruption in the users spatial memory, particularly if the window is significantly resized. However, in their work characterising how windows are used, Gaylin [12] observed that windows were generally resized when users initially logged on (generally to make them larger) rather than later during their work. One possibility to resolve this issue would be to always maintain the thumbnails in the same position and simply resize them appropriately to fit inside the window. However, this would mean they no longer met the "space-filling" criterion. This is an allowable trade-off for gains resulting from maintaining the thumbnails in their original positions.

5.2.5 Document Editing

This is a similar issue to that discussed in section 5.2.4, except when editing a document one may be adding, deleting or even moving content between pages. These activities are likely to disrupt a users spatial memory. However, as was seen in the second set of tasks for experiment three, users were still able to locate pages that had been moved quicker than a pure visual search. If pages are added to the end of a document, all previous pages will maintain their positions in the matrix of thumbnails relative to those around them.

5.3 Future Work

The time limitations of this work meant that many related areas of research could not be pursued. Described in this section are areas that would be interesting to consider in future work.

5.3.1 Spatial Memory Ability

The participant samples used in these experiments would have encompassed persons with both strong and weak spatial memory abilities. In these evaluations we did not discriminate between these two groups of people. It would make interesting future work to evaluate participants spatial abilities before conducting evaluation tasks with the document navigation interfaces. We would expect people with strong spatial abilities to perform better with the Space-Filling Thumbnails interface than those with weak spatial relocation abilities. It would be of interest to determine whether those with weak spatial memory received the same level of benefits for page re-acquisition tasks with the SFT interface.

5.3.2 Semantic Thumbnail Zooming

Presently the thumbnails that are displayed are simply scaled down versions of the original pages. To aid in search tasks semantic zooming could be added. This would highlight interesting features on a page, such as headings, diagrams or page numbers. One technique described by Woodruff et al. [36] combines textual summaries and thumbnails to produce “enhanced thumbnails”. Although applied to navigation between documents (web pages) the same concept could easily be applied to navigation within documents.

5.3.3 Find Facilities

In this research we have concentrated on visual and spatial searches. However, visual search is not always the optimal technique for searching for objects in a document. Often *Find* functionalities of a system are used to identify “interesting” areas of a document (for instance a search for “subjective” in this document). It would make interesting future work in implementing a Find facility for the Space-Filling Thumbnails interface. As the SFT interface provides a complete overview of a document, one could envision that pages that match the search query are highlighted. This could be further enhanced by changing the shade of highlighting depending on the number of matches in a page. This would quickly provide the user with an indication of the “interesting” areas of the document—occurrences of the search terms in the contents or index would be able to be easily skipped.

5.3.4 Modal Switching

One concern with the SFT interface is that it is a modal system. The key thing that any modal system *must* provide is the ability to quickly distinguish which mode the system is in. With the SFT system this is clear from the visualisation on screen. Users can easily determine whether they are in thumbnail mode or full zoom mode. The only discrepancy may arise when one page is being viewed in the SFT system, in which case the thumbnail view will be disabled. Hornbæk et al. [17] in their evaluation of interfaces with and without overviews, found that switching between an overview and a detail view required large mental effort and time on the part of the user. We believe that this is not the case with the SFT system, as switching between modes is eased by animation. However, it would be interesting to see if there is a higher cognitive load on the user when this action takes place than in a traditional scrollbar based system. One can argue that there is a high mental demand in switching from using the thumbnails to reading a page using the Thumbnail-Enhanced Scrollbar system.

5.3.5 Hick-Hyman Law for Choice Reaction

The Hick-Hyman law for choice reaction describes the time taken for a user to make a decision given a finite number of choices. If the SFT interface were to perform worse than that predicted by this law, one should be concerned and investigation would be necessary to determine what was impeding the interface from conforming to this law. Unfortunately we do not have enough data points to be able to confidently do this analysis. This would require evaluating the interface using many different document lengths. This would make for interesting future work.

5.3.6 Power Law of Practise

The Power Law of Practise says that the more we practise a task the quicker we become at performing that task. Most learning occurs at the beginning of the practise. One would expect that participants would learn the location of the pages being searched for according to this law. To test this hypothesis an experiment would need to be run where participants searched for the same page many times (up to ten to provide an accurate analysis). Again if SFT did not perform according to this law, an investigation would be necessary to determine the factors impeding the learning.

5.3.7 Non-document Applications

In this study we have concentrated on evaluating Space-Filling Thumbnails for *document navigation*. However, this system has the potential to be applied to navigation with any form of object that cannot fit satisfactorily on a screen. The same concepts apply to images, maps and diagrams, the only concession is that these objects would require discrete divisions to allow zooming to take place as in SFT. For instance in a class diagram, divisions may be made by class, so that the user can easily inspect one class and then zoom out and see it in its context again.

6 Conclusions

We have designed, implemented and evaluated the Space-Filling Thumbnails document navigation interface. This system has shown to have faster visual search and spatial re-acquisition capabilities than all other systems it was evaluated against.

Firstly, Space-Filling Thumbnails was compared with other previously proposed navigation systems—both mainstream and research—including traditional scrollbars. The SFT system was the fastest for both visual and spatial search tasks. The interface was then evaluated against the “best of the rest,” Thumbnail-Enhanced Scrollbars, in a series of tests with a variety of document lengths and implicitly types. In this case the spatial location of pages was not artificially implanted into participants memory. Again, the Space-Filling Thumbnails interface was faster for all tasks. In a final experiment we evaluated the possible weakness of SFT—large documents with no distinct features. Despite the small thumbnails generated from a 300 page book, SFT still had faster search times than its competitor TES. In tasks where pages were moved to imitate document editing, SFT came out on top.

Subjective evaluations of the Space-Filling Thumbnails interface gathered throughout these experiments indicate that it would be the preferred interface for these sorts of tasks.

In this study we have shown the Space-Filling Thumbnails interface to be a viable addition, if not replacement, for the traditional scrollbar. We believe its fast visual search and exploitation of spatial memory mean it will play a pivotal role in document navigation systems of the future.

Bibliography

[1] BEDERSON, B. B. Photomesa: A Zoomable Image Browser Using Quantum Treemaps And Bubblemaps. In *UIST '01: Proceedings of the 14th Annual ACM Symposium on User Interface Software and Technology* (New York, NY, USA, 2001), ACM Press, pp. 71–80.

[2] BJÖRK, S. Hierarchical Flip Zooming: Enabling Parallel Exploration Of Hierarchical Visualizations. In *AVI '00: Proceedings of the Working Conference on Advanced Visual Interfaces* (2000), ACM Press, pp. 232–237.

[3] BYRD, D. A Scrollbar-based Visualization for Document Navigation. In *Proceedings of Digital Libraries '99* Berkeley CA, August (1999), pp. 122–129.

[4] BYRNE, M., JOHN, B., WEHRLE, N., AND CROW, D. The Tangled Web We Wove: A Taskonomy of WWW Use. In *Proceedings of CHI'99 Conference on Human Factors in Computing Systems* Pittsburgh, May 15–20 (1999), pp. 544–551.

[5] COCKBURN, A., SAVAGE, J., AND WALLACE, A. Tuning and Testing Scrolling Interfaces that Automatically Zoom. In *CHI '05: Proceeding of the SIGCHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2005), ACM Press, pp. 71–80.

[6] COCKBURN, A AND SAVAGE, J. Comparing Speed-Dependent Automatic Zooming with Traditional Scroll, Pan and Zoom Methods. In *People and Computers XVII (Proceedings of the 2003 British Computer Society Conference on Human-Computer Interaction)* (2003), pp. 87–102.

[7] CZERWINSKI, M. P., VAN DANTZICH, M., ROBERTSON, G., AND HOFFMAN, H. The Contribution of Thumbnail Image, Mouse-over Text and Spatial Location Memory to Web Page Retrieval in 3D. In *Proceedings of IFIP INTERACT'99: Human-Computer Interaction* (1999), Viewing the Web, pp. 163–170.

[8] DE BRUIJN, O., AND SPENCE, R. Rapid Serial Visual Presentation, 2000. Available from: `http://www.iis.ee.ic.ac.uk/~o.debruijn/rsvp.pdf`.

[9] DUMAIS, S. T., AND JONES, W. P. A Comparison of Symbolic and Spatial Filing. In *CHI '85: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 1985), ACM Press, pp. 127–130.

[10] EHRET, B. Learning Where to Look: Location Learning in Graphical User Interfaces. In *Proceedings of CHI'2002 Conference on Human Factors in Computing Systems. CHI Letters 4(1).* Minneapolis, Minnesota, 20–25 April (2002), pp. 211–218.

[11] FITTS, P. M. The Information Capacity of the Human Motor System in Controlling the Amplitude of Movement. *Journal of Experimental Psychology 47*, 6 (1954), 381 – 391.

[12] GAYLIN, K. B. How are Windows Used? Some Notes on Creating an Empirically-Based Windowing Benchmark Task. In *CHI '86: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 1986), ACM Press, pp. 96–100.

[13] GINSBURG, A., MARKS, J., AND SHIEBER, S. A viewer for PostScript Documents. In *UIST '96: Proceedings of the 9th annual ACM Symposium on User Interface Software and Technology* (New York, NY, USA, 1996), ACM Press, pp. 31–32.

[14] HART, S., AND STAVELAND, L. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Human Mental Workload*, P. Hancock and N. Meshkati, Eds. Elsevier Science, 1988, pp. 139–183.

[15] HOEBEN, A., AND STAPPERS, P. J. Flicking Through Page-Based Documents with Thumbnail Sliders and Electronic Dog-Ears. In *CHI '00: CHI '00 Extended Abstracts on Human Factors in Computing Systems* (New York, NY, USA, 2000), ACM Press, pp. 191–192.

[16] HORNBÆK, K AND FRØKJÆR, E. Reading of Electronic Documents: The Usability of Linear, Fisheye, and Overview+Detail Interfaces. In *Proceedings of CHI'2001 Conference on Human Factors in Computing Systems. CHI Letters 3(1).* Seattle, Washington, March 31–April 6 (2001), pp. 293–300.

[17] HORNBAEK, K., BEDERSON, B., AND PLAISANT, C. Navigation Patterns and Usability of Zoomable User Interfaces With and Without an Overview. *ACM Transactions on Computer-Human Interaction 9*, 4 (2002), 362–389.

[18] HYMAN, R. Stimulus Information as a Determinant of Reaction Time. *Journal of Experimental Psychology 45* (1953), 188–196.

[19] IGARASHI, T., AND HINCKLEY, K. Speed-Dependent Automatic Zooming for Browsing Large Documents. In *UIST '00: Proceedings of the 13th annual ACM symposium on User interface software and technology* (New York, NY, USA, 2000), ACM Press, pp. 139–148.

[20] INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. ISO 216:1975, Writing Paper and Certain Classes of Printed Matter — Trimmed sizes — A and B series, 1975. Available from: `http://www.iso.org/`.

[21] JONES, W. P., AND DUMAIS, S. T. The Spatial Metaphor for User Interfaces: Experimental Tests of Reference By Location Versus Name. *ACM Trans. Inf. Syst. 4*, 1 (1986), 42–63.

[22] LAAKSO, S., LAAKSO, K., AND SAURA, A. Improved Scroll Bars. In *Extended Abstract of CHI'2000. ACM Conference on Human Factors in Computing Systems* The Hague, The Netherlands (2000), ACM Press, pp. 97–98.

[23] LANDAUER, T. K., AND NACHBAR, D. W. Selection from Alphabetic and Numeric Menu Trees Using a Touch Screen: Breadth, Depth, and Width. In *CHI '85: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 1985), ACM Press, pp. 73–78.

[24] MANDLER, J., SEEGMILLER, D., AND DAY, J. On the Encoding of Spatial Information. *Memory Cognition 5* (1977), 10–16.

[25] MCCRICKARD, D. S., AND CATRAMBONE, R. Beyond the scrollbar: An evolution and evaluation of alternative navigation techniques. In *Visual Languages* (1999), pp. 270–. Available from: `citeseer.ist.psu.edu/article/mccrickard99beyond.html`.

[26] MURRAY, K. B., AND HÄUBL, G. A Human Capital Perspective of Skill Acquisition and Interface Loyalty. *Commun. ACM 46*, 12 (2003), 272–278.

[27] NEWELL, A., AND ROSENBLOOM, P. S. Mechanisms of Skill Acquisition and the Law of Practice. *Cognitive Skills and Their Acquisition* (1981), 1–55.

[28] O'HARA, K., AND SELLEN, A. A Comparison of Reading Paper and On-Line Documents. In *CHI '97: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 1997), ACM Press, pp. 335–342.

[29] O'HARA, K., SELLEN, A., AND BENTLEY, R. Supporting Memory for Spatial Location while Reading from Small Displays. In *CHI '99: CHI '99 Extended Abstracts on Human Factors in Computing Systems* (New York, NY, USA, 1999), ACM Press, pp. 220–221.

[30] ROBERTSON, G., CZERWINSKI, M., LARSON, K., ROBBINS, D. C., THIEL, D., AND VAN DANTZICH, M. Data Mountain: Using Spatial Memory for Document Management. In *UIST '98: Proceedings of the 11th annual ACM Symposium on User Interface Software and Technology* (New York, NY, USA, 1998), ACM Press, pp. 153–162.

[31] ROBERTSON, G. G., AND MACKINLAY, J. D. The Document Lens. In *UIST '93: Proceedings of the 6th Annual ACM Symposium on User Interface Software and Technology* (1993), ACM Press, pp. 101–108.

[32] SAVAGE, J., AND COCKBURN, A. Comparing Automatic and Manual Zooming Methods for Acquiring Off-Screen Targets. In *People and Computers XIX (Proceedings of the 2005 British Computer Society Conference on Human-Computer Interaction)*, Edinburgh, UK (2005), Springer-Verlag.

[33] SMITH, G. M., AND SCHRAEFEL, M. C. The Radial Scroll Tool: Scrolling Support for Stylus- or Touch-Based Document Navigation. In *UIST '04: Proceedings of the 17th Annual ACM Symposium on User Interface Software and Technology* (New York, NY, USA, 2004), ACM Press, pp. 53–56.

[34] SPENCE, B., WITKOWSKI, M., FAWCETT, C., CRAFT, B., AND DE BRUIJN, O. Image Presentation in Space and Time: Errors, Preferences and Eye-Gaze Activity. In *AVI '04: Proceedings of the Working Conference on Advanced Visual Interfaces* (2004), ACM Press, pp. 141–149.

[35] SUN, L., AND GUIMBRETIÈRE, F. Flipper: A New Method of Digital Document Navigation. In *CHI '05: CHI '05 Extended Abstracts on Human Factors in Computing Systems* (New York, NY, USA, 2005), ACM Press, pp. 2001–2004.

[36] WOODRUFF, A., FAULRING, A., ROSENHOLTZ, R., MORRSION, J., AND PIROLLI, P. Using Thumbnails to Search the Web. In *CHI '01: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2001), ACM Press, pp. 198–205.

A Evaluation Documents

The documents listed here were used for the evaluations conducted during this research.

Experiment One

Training Document

- BICKMORE, T. W., AND PICARD, R. W. Establishing and Maintaining Long-Term Human-Computer Relationships. *ACM Trans on Computer-Human Interaction 12*, 2 (2005), 293–327.

Evaluation Documents

- BELLINI, P., NESI, P., AND SPINU, M. B. Cooperative Visual Manipulation of Music Notation. *ACM Trans on Computer-Human Interaction 9*, 3 (2002), 194–237.

- BROTHERTON, J. A., AND ABOWD, G. D. Lessons Learned From eClass: Assessing Automated Capture and Access in the Classroom. *ACM Trans on Computer-Human Interaction 11*, 2 (2004), 121–155.

- BURNETT, M., YANG, S., AND SUMMET, J. A Scalable Method for Deductive Generalization in the Spreadsheet Paradigm. *ACM Trans on Computer-Human Interaction 9*, 4 (2002), 253–284.

- K, K. H., AND KJÆR, E. F. Reading Patterns and Usability in Visualizations Of Electronic Documents. *ACM Trans on Computer-Human Interaction 10*, 2 (2003), 119–149.

- PETERSEN, M. G., MADSEN, K. H., AND R, A. K. The Usability of Everyday Technology: Emerging and Fading Opportunities. *ACM Trans on Computer-Human Interaction 9*, 2 (2002), 74–105.

- ULLMER, B., ISHII, H., AND JACOB, R. J. K. Token+Constraint Systems for Tangible Interaction with Digital Information. *ACM Trans on Computer-Human Interaction 12*, 1 (2005), 81–118.

- VELEZ, M., TREMAINE, M. M., SARCEVIC, A., DOROHONCEANU, B., KREBS, A., AND MARSIC, I. ”Who’s in Charge Here?” Communicating Across Unequal Computer Platforms. *ACM Trans on Computer-Human Interaction 11*, 4 (2004), 407–444.

Experiment Two

Training Document

- T. W. Bickmore and R. W. Picard, “Establishing and Maintaining Long-Term Human-Computer Relationships,” *ACM Trans. Comput.-Hum. Interact.*, vol. 12, no. 2, pp. 293–327, 2005.

10 Page Documents

- AHLSTRÖM, D. Modeling and Improving Selection in Cascading Pull-Down Menus Using Fitts’ Law, the Steering Law and Force Fields. In *CHI ’05: Proceedings of the SIGCHI conference on Human Factors in Computing Systems* (New York, NY, USA, 2005), ACM Press, pp. 61–70.

- MARSHALL, C. C., AND BLY, S. Saving and Using Encountered Information: Implications for Electronic Periodicals. In *CHI '05: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2005), ACM Press, pp. 111–120.

30 Page Documents

- MENG, H., CHING, P. C., CHAN, S. F., WONG, Y. F., AND CHAN, C. C. Isis: An Adaptive, Trilingual Conversational System with Interleaving Interaction and Delegation Dialogs. *ACM Transactions on Computer-Human Interaction 11*, 3 (2004), 268–299.

- VELEZ, M., TREMAINE, M. M., SARCEVIC, A., DOROHONCEANU, B., KREBS, A., AND MARSIC, I. "Who's in Charge Here?" Communicating Across Unequal Computer Platforms. *ACM Transactions on Computer-Human Interaction 11*, 4 (2004), 407–444.

150 Page Documents

- SONY CORPORATION. Digital Videocassette Recorder Operation Instructions, DVCAM DSR-2000/2000P. Instruction Manual. Available from: `http://www.sony.ca/dvcam/manuals.htm`.

- THE GLOBAL OLYMPUS GROUP. E-300 Digital Camera Advanced Manual. Instruction Manual. Available from: `http://www.olympus.co.jp/en/support/imsg/digicamera/download/manual/esystem/man_e300_en.pdf`.

Experiment Three and Four

- TOLSTOY, L. *Anna Karenina*. Bookstacks, 2004. Electronic Version from `www.bookstacks.org`.